

# Fakultät für Physik und Astronomie

## Ruprecht-Karls-Universität Heidelberg

Diplomarbeit  
im Studiengang Physik

vorgelegt von  
**Peter Schneider**  
aus Neckarsulm

1998



**Simulation und Visualisierung  
elektrischer und optischer Eigenschaften  
von Halbleiterbauelementen**

Die Diplomarbeit wurde von Peter Schneider ausgeführt am  
Institut für Hochenergiephysik der Universität Heidelberg  
unter der Betreuung von  
Herrn Prof. Dr. K. Meier



## **Inhalt**

Bauelement-Simulatoren ermöglichen Einblicke in die inneren physikalischen Funktionen von Elementen der Halbleitertechnik. In der vorliegenden Arbeit werden mit Hilfe eines kommerziellen Software-Systems die elektrischen und optischen Eigenschaften von Halbleiterbauelementen untersucht, sowie deren elektrisches Verhalten bei Modifikationen der Geometrie und Störstellenverteilung evaluiert. Die Architekturen der Elemente werden anhand von Dopingprofilen des  $0,8\mu m$  CMOS/BiCMOS-Prozesses einer Chip-Herstellerfirma erstellt, mit der das ASIC-Labor Heidelberg zusammenarbeitet. Dabei werden sowohl die intrinsischen, als auch die parasitären Eigenschaften von Bausteinen der Mikroelektronik beschrieben. Die Ergebnisse werden mit Hilfe farblich abgestufter Konturplots visualisiert.

## **Abstract**

Device-Simulators offer the possibility to gain an insight into the physical functions of semiconductor devices. This thesis shows the electrical and optical characteristics of microelectronic devices using a commercial software system. Additionally it determines the electrical behaviour while modifying geometry and density of impurities. In order to layout the devices rather realistically, they are designed with doping profiles of the  $0,8\mu m$  CMOS/BiCMOS-process of a chip manufacturer, which cooperates with the ASIC-Laboratory Heidelberg. Intrinsic as well as parasitic properties of semiconductor devices are also evaluated. The results will be visualized in forms of rainbow coloured contour plots.



# Inhaltsverzeichnis

<b>Einführung</b>	<b>1</b>
<b>1 Physikalische Grundlagen</b>	<b>3</b>
1.1 Bandstruktur und Energiebänder . . . . .	3
1.2 Intrinsische Halbleiter . . . . .	6
1.2.1 Direkte und indirekte Halbleiter . . . . .	7
1.2.2 Ladungsträgerkonzentration . . . . .	8
1.3 Dotierte Halbleiter . . . . .	9
1.3.1 Ladungsträgerkonzentration . . . . .	12
<b>2 Numerische Methoden</b>	<b>15</b>
2.1 Grundgleichungen . . . . .	16
2.2 Sekundäre Gleichungen . . . . .	17
2.3 Beweglichkeit . . . . .	24
2.4 Generation und Rekombination von Ladungsträgern . . . . .	31
<b>3 Der pn-Übergang</b>	<b>35</b>
3.1 Symmetrisch abrupt dotierte Diode . . . . .	36
3.2 N-Diffusion/Substrat Diode . . . . .	41
3.3 N-Wanne/Substrat Diode . . . . .	45
3.4 P-Diffusion/N-Wanne Diode . . . . .	47
<b>4 Bipolar Transistoren</b>	<b>49</b>
4.1 Der Parasitäre Vertikale PNP-Transistor . . . . .	50
4.2 Der BiCMOS NPN-Transistor . . . . .	56
<b>5 MOS-Transistor</b>	<b>65</b>
5.1 Der NMOS-Transistor . . . . .	65
<b>6 Optische Simulationen</b>	<b>73</b>
6.1 Photostrom in Abhängigkeit der angelegten Spannungen . . . . .	75
6.1.1 N-Diffusion/Substrat Diode . . . . .	75
6.1.2 N-Wanne/Substrat Diode . . . . .	77
6.1.3 P-Diffusion/N-Wanne Diode . . . . .	78
6.1.4 Parasitärer Vertikaler PNP-Transistor . . . . .	80
6.1.5 BiCMOS NPN-Transistor . . . . .	82
6.2 Räumliche Antwort . . . . .	84

6.2.1	N-Diffusion/Substrat Diode . . . . .	84
6.2.2	N-Wanne/Substrat Diode . . . . .	84
6.2.3	P-Diffusion/N-Wanne Diode . . . . .	87
6.2.4	Parasitärer Vertikaler PNP-Transistor . . . . .	89
6.2.5	BiCMOS NPN-Transistor . . . . .	90
6.2.6	NMOS-Transistor . . . . .	91
6.3	Spektrale Antwort . . . . .	92
6.3.1	N-Diffusion/Substrat Diode . . . . .	92
6.3.2	N-Wanne/Substrat Diode . . . . .	93
6.3.3	P-Diffusion/N-Wanne Diode . . . . .	94
6.3.4	Parasitärer Vertikaler PNP-Transistor . . . . .	95
6.3.5	BiCMOS NPN-Transistor . . . . .	96
6.3.6	NMOS Transistor . . . . .	97
6.4	Transiente Analyse . . . . .	98
6.4.1	N-Diffusion/Substrate Diode . . . . .	98
6.4.2	N-Wanne/Substrat Diode . . . . .	98
6.4.3	P-Diffusion/N-Wanne Diode . . . . .	99
6.4.4	Parasitärer Vertikaler PNP-Transistor . . . . .	100
6.4.5	BiCMOS NPN-Transistor . . . . .	101
6.5	Guard . . . . .	102
<b>7</b>	<b>Messungen</b>	<b>107</b>
7.1	Bipolar . . . . .	107
7.2	NMOS . . . . .	109
	<b>Zusammenfassung</b>	<b>113</b>
	<b>A Maßstabsgetreue Abbildungen</b>	<b>117</b>
	<b>B Elektronenzustände im Festkörper</b>	<b>125</b>
B.1	Freies Elektronengas . . . . .	125
B.2	Periodisches Potential . . . . .	128
B.3	Effektive Masse . . . . .	130
	<b>C Deckbuild-Datei</b>	<b>132</b>
	<b>D Skript zur Animation</b>	<b>134</b>
	<b>Literaturverzeichnis</b>	<b>137</b>



# Einführung

Halbleiter sind aus Forschung und Technik nicht mehr wegzudenken. Dabei werden sie meist als Werkzeug, vor allem in Form von Rechenanlagen, benutzt, um andere Bereiche schneller, genauer und besser erschließen zu können. Man denke zum Beispiel an die Datenflut, die in Experimenten der Hochenergiephysik anfällt, und deren Auswertung, ohne Zuhilfenahme moderner und leistungsfähiger Computer, mit nicht überschaubarem Zeitaufwand verbunden wäre.

Aber auch im privaten Bereich können sich die meisten Menschen der industrialisierten Länder ein Leben ohne Halbleiter kaum mehr vorstellen. Es seien hier nur Telekommunikation und Unterhaltungselektronik genannt. Immer schnellere und leistungsfähigere Halbleiter-Chips haben das gesamte Leben der Industrienationen verändert.

Im ASIC-Labor<sup>1</sup> der Universität Heidelberg werden solche miniaturisierten Schaltungen für verschiedene spezielle Anwendungen entwickelt. Ein Projekt zum Beispiel, befaßt sich mit der Entwicklung einer Orientierungshilfe für sehbehinderte oder blinde Menschen. Dazu werden auf *einem* Chip sowohl die optische Erfassung der Umgebung mit Halbleiter-Photosensoren, als auch die elektronische Schaltung zur Aufbereitung der Signale realisiert. Das hat einerseits Vorteile, da die Wege, die die Ströme zurücklegen müssen klein sind, und daher wenig Verluste auftreten und die hohe Integration der Bauteile die Sehhilfe portabel und handlich gestaltet. Andererseits trifft nun das Licht nicht nur auf die optischen Sensoren, sondern auch auf die, sich in unmittelbarer Nähe befindende, Auswerteelektronik. Da sich diese aber in der physikalischen Wirkungsweise nicht von den Photosensoren unterscheidet, wird ihr Verhalten, wenn auch gering, vom einfallenden Licht ebenfalls beeinflusst. Außerdem beeinflussen sich benachbarte Elemente durch Diffusion der Ladungsträger gegenseitig, so daß ihr elektrisches Verhalten wechselseitig modifiziert wird. In der vorliegenden Arbeit soll daher versucht werden einen Einblick in die inneren Vorgänge der Halbleiterbauelemente zu bekommen. Dazu dient als Werkzeug ein Bauelement-Simulator. Dem Simulator werden die Geometrie und die Verteilung der Störstellendichten (Dotierprofile) des Halbleiterbauelementes eingegeben. Weiterhin werden die Eigenschaften der verwendeten Materialien, wie zum Beispiel der energetische Abstand der Bandlücke etc., angegeben. Schließlich werden Anschlußkontakte definiert und mit elektrischen Potentialen versehen. Der Simulator berechnet damit die elektrischen Eigenschaften und bietet die Möglichkeit diese in verschiedener Weise darzustellen. Um eine möglichst gute Übereinstimmung mit den im ASIC-Labor Heidelberg verwendeten Strukturen zu erzielen, wurden die Dotierprofile von real existierenden Chips verwendet. Dazu wurden

---

<sup>1</sup>ASIC: Application Specific Integrated Circuit

von der Firma *Austria Mikro Systeme International AG* dem ASIC-Labor Heidelberg freundlicherweise sowohl die Dotierprofile ihres  $0,8\mu\text{m}$  CMOS-Prozesses, als auch diejenigen des BiCMOS-Prozesses übermittelt.

Anhand dieser Daten wurden verschiedene Bipolar-Strukturen (Dioden und Transistoren) und ein NMOS-Transistor erstellt.

In den Kapiteln 3 bis 5 werden diese Strukturen auf ihr elektrisches Verhalten untersucht und dieses unter anderem in Form von farblich abgestuften Regenbogenskalen dargestellt. Der Simulator bietet außerdem die Möglichkeit, die Bauelemente mit elektromagnetischen Wellen im sichtbaren Bereich zu bestrahlen, um so deren optischen Eigenschaften zu bestimmen. Im Kapitel 6 werden die Bauelemente unter optischem Einfluß simuliert und die Resultate ebenfalls mit den erwähnten Methoden dargestellt. Zum Schluß werden die simulierten Ergebnisse mit Messungen verglichen.

Zunächst sollen aber im folgenden Kapitel kurz die physikalischen Grundlagen der Halbleiterphysik zusammengestellt werden, die als Gedächtnisstütze gedacht sind und insofern nichts Neues enthalten.

# Kapitel 1

## Physikalische Grundlagen

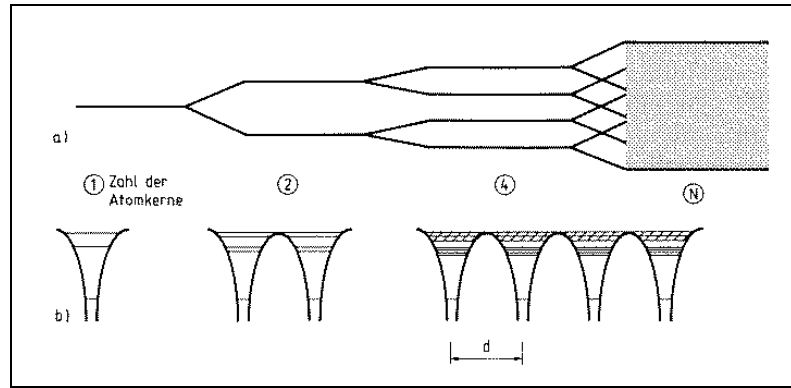
### 1.1 Bandstruktur und Energiebänder

Isolierte Atome besitzen diskrete Energieniveaus. Verringert man den Abstand isolierter Atome derart, daß sie einen Festkörper bilden (d.h. sie gehen eine chemische Bindung ein), so werden die möglichen Zustände der vormals isolierten Atome durch das Potential der benachbarten Atomkerne mehr oder weniger stark beeinflusst. Diese Beeinflussung spüren vor allem die Elektronen der äußersten Schale, die *Valenzelektronen*, während die inneren, sogenannte *Rumpfelektronen*, fast ausschließlich das Potential ihres eigenen Atomkernes erfahren. Durch das von den Nachbaratomen erzeugte Potential spalten die diskreten Energieniveaus der Valenzelektronen in viele, dicht benachbarte Energieniveaus auf. Da ein makroskopischer Festkörper in der Größenordnung  $10^{23}$  Atome besitzt, liegen diese Energieniveaus derart dicht beieinander, daß man von einem quasikontinuierlichen „Band“ sprechen kann. Besteht ein Kristall aus  $N$  Elementarzellen mit je einem Valenzelektron pro Zelle, dann werden im entsprechenden Band des Festkörpers  $2N$  Energiezustände entstehen<sup>1</sup>, die von Elektronen besetzt werden können. In Abbildung 1.1 (a) ist zu sehen, wie die diskreten Energieniveaus eines isolierten Atomes durch Hinzufügen weiterer Atome in sich letztendlich überlappende Energieniveaus übergehen. Der untere Teil der Abbildung 1.1 soll illustrieren, daß sich die tieferen Niveaus nicht so stark von den Nachbarpotentialen beeinflussen lassen und somit nur eine geringe Aufspaltung erfahren ( $d$  bezeichnet den Abstand der Atomkerne).

Die Aufspaltung der Niveaus hängt von der Distanz der benachbarten Atome ab, aber nicht von deren Zahl, das heißt die Niveaus fallen – bei äquidistanter Anordnung der Atome – alle in etwa denselben Energiebereich. Das bedeutet, daß ein Elektron innerhalb eines Bandes quasikontinuierliche Energiezustände einnehmen kann, was unmittelbar Konsequenzen für die elektrische Leitfähigkeit hat. Die Aufspaltung der ursprünglich diskreten Energieniveaus der Atome geschieht nun mehr oder weniger stark (je nach Abstand zum Kern) mit allen Energieniveaus, so daß sich im Festkörper mehrere Bänder ausbilden. Die Gesamtheit aller Energiewerte, die Elektronen im Festkörper einnehmen können bezeichnet man als *Bandstruktur*. Die Bandstruktur entspricht also im Festkörper dem Termschema des freien Atoms. Die Energieniveaus werden nun sukzessive von unten mit Elektronen besetzt, so wie dies auch

---

<sup>1</sup>Jeder Term des isolierten Atoms spaltet in  $N$  Energierme auf, diese können aber aufgrund des Spins von zwei Elektronen besetzt werden.



**Abbildung 1.1:** Entstehung von Energiebändern durch Aneinanderreihung von Atomen (aus [1])

beim freien Atom geschieht. Am absoluten Nullpunkt der Temperatur nehmen die Elektronen die niedrigsten zur Verfügung stehenden Energiezustände ein. Bei endlichen Temperaturen muß die Besetzung durch eine Verteilungsfunktion geregelt werden. Weil Elektronen *Fermionen* sind, genügen sie der *Fermi-Dirac* Statistik. Am absoluten Nullpunkt ist die fermi'sche Verteilungsfunktion eine Stufenfunktion, die an der Stelle  $E = E_F$ <sup>2</sup> von 0 auf 1 springt. In der Abbildung B.2 im Anhang ist die Fermi-Dirac Funktion für zwei endliche Temperaturen gezeigt. Es ist zu erkennen, daß sich die Besetzung der Zustände bei üblichen Temperaturen (insbesondere Zimmertemperatur) nur wenig von derjenigen am absoluten Nullpunkt unterscheidet. Die Besetzung der Zustände erfolgt also bei Zimmertemperatur nahezu so wie am absoluten Nullpunkt der Temperatur. Lediglich der Rand der *Fermikugel*<sup>3</sup> wird ein wenig „aufgeweicht“. Durch die Bildung von Bändern sind die Elektronen nicht mehr am Gitteratom lokalisiert, sondern können sich innerhalb des Bandes bewegen. Der Energiezustand des Elektrons ist also eine Eigenschaft des kompletten makroskopischen Festkörpers.

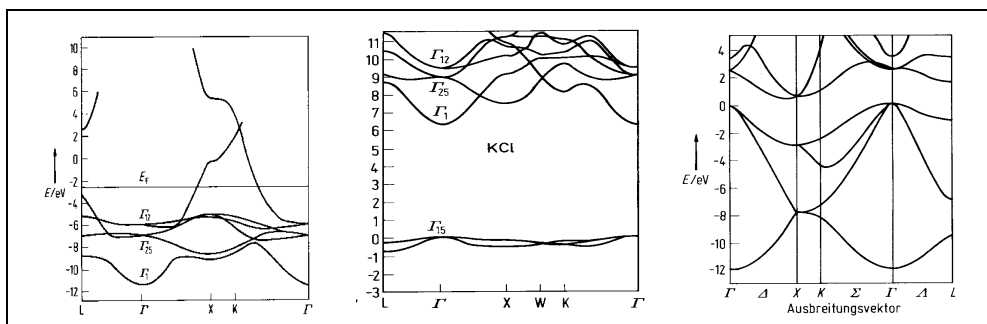
Aufgrund der Bildung von Bändern, der damit verbundenen Delokalisierung der Elektronen von den Gitteratomen und der Existenz quasikontinuierlicher Energiezustände innerhalb eines Bandes läßt sich das Verhalten von Festkörpern bezüglich ihrer elektronischen Leitfähigkeit verstehen. Dazu muß man aber wissen, wie die Energieniveaus der Bänder mit Elektronen besetzt sind. Ein Band, das zum Beispiel aus einer vollbesetzten Unterschale eines Atoms hervorgeht, wird auch im Festkörper „voll besetzt“ sein, das heißt alle möglichen Energiezustände in diesem Band werden durch Elektronen eingenommen. Dieses Band kann dann nicht zur elektrischen Leitfähigkeit beitragen, denn ein Elektron kann sich nur dann innerhalb eines Bandes „bewegen“, wenn darin noch freie Zustände existieren, in die es etwa aufgrund äußerer Energiezufuhr (z.B. externes elektrisches Feld) *gestreut* werden kann<sup>4</sup>. Existieren in einem Band aber mehr Zustände als Elektronen, die diese besetzen können, so ist es nur teilweise besetzt, und ein Elektron findet innerhalb dieses Bandes freie Zustände, in die es gestreut werden kann (Stromfluß). Solch ein Band entsteht zum Beispiel aus einem Energie-

<sup>2</sup> $E_F$  bezeichnet die *Fermienergie*, die in Anhang A für das freie Elektronengas hergeleitet wird

<sup>3</sup>Stellt man die erlaubten Zustände freier Elektronen im Kastenpotential im  $\vec{k}$ -Raum dar, so ergeben sich als Flächen konstanter Energie Kugeln ( $E = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2)$ )

<sup>4</sup>In vollbesetzten Bändern kann zwar eine *Umbesetzung* der Elektronen stattfinden, diese liefert aber keinen Beitrag zum Netto-Stromfluß.

niveau einer nur teilweise besetzten äußeren Schale des isolierten Atoms. Man bezeichnet das höchste mit Elektronen voll besetzte Band als *Valenzband* und das niedrigste teilweise besetzte Band als *Leitungsband*. Die Bänder können sich im Festkörper teilweise überlappen oder sie sind durch Lücken voneinander getrennt, den *verbotenen Zonen*. Die verbotenen Zonen sind Energiebereiche, in denen sich Elektronen nicht aufhalten können. Wenn die Energielücke zwischen Valenzband und Leitungsband nicht so groß ist, daß Elektronen vom Valenzband ins Leitungsband gelangen können, so ist dadurch ebenfalls ein Stromfluß möglich, da ein angehobenes Elektron im leeren Leitungsband natürlich freie Zustände vorfindet, in die es gestreut werden kann. Als grobe Regel kann man sagen, daß Atome mit ungerader Anzahl von Valenzelektronen zu elektrischen Leitern werden, wogegen solche Atome mit gerader Anzahl von Valenzelektronen sich zu Festkörpern mit vollbesetzten Energiebändern, und damit zu Isolatoren bzw. Halbleitern, ausbilden.



**Abbildung 1.2:** Berechnete Bandstrukturen eines Metalls (nach Segall), eines Isolators (nach de Cicco) und eines Halbleiters (nach Yin und Cohen). Die Bezeichnungen der Abszisse geben die Richtungen in der Brioullin-Zone an (aus [4])

Wie unterscheiden sich Metalle, Halbleiter und Isolatoren in ihrer Bandstruktur?

Die Abbildung 1.2 zeigt Bandstrukturen dreier sich elektrisch unterschiedlich verhaltender Materialien. Im linken Bild, das die Bandstruktur eines Metalles darstellt, ist zusätzlich die Fermienergie  $E_F$  eingezeichnet, die sich hier *innerhalb* eines Bandes befindet, weshalb dieses nicht voll besetzt ist und somit zur elektrischen Leitfähigkeit beiträgt. Im mittleren ist ein typischer Isolator zu sehen, der eine Bandlücke von mehreren  $eV$  aufweist. Das Fermi-Niveau, das hier nicht eingezeichnet ist, liegt in der verbotenen Zone, so daß das Leitungsband leer und das Valenzband voll besetzt ist. Ganz rechts ist die berechnete Bandstruktur von Silizium gezeigt. Die Bandlücke ist aber in Wirklichkeit größer (ca.  $1.1eV$ ) als hier dargestellt. Um die elektrischen Eigenschaften zu verstehen, ist die genaue Struktur der Bänder, also die Beziehung zwischen Wellenzahlvektor  $\vec{k}$  und der dazugehörigen Energie  $E$ , nicht so von Bedeutung, so daß man zu einem vereinfachten Modell übergeht, das nur die unterste Kante des Leitungsbandes und die oberste des Valenzbandes betrachtet. Dieses Modell wird als *Bänderschema* bezeichnet. Bei Metallen gibt es grundsätzlich zwei verschiedene Möglichkeiten, wie elektrische Leitfähigkeit zustande kommen kann. Im ersten Fall liegt die Fermienergie innerhalb des Leitungsbandes, so daß dieses nur teilweise besetzt ist und so Elektronen durch kontinuierliche Energiezufuhr<sup>5</sup> in freie Zustände gestreut werden können. Im zweiten Fall überlappen

<sup>5</sup>Im Prinzip sind die Energiezustände diskret, sie liegen aber so dicht, daß man von einem quasikontinuierlichem Band sprechen kann.

sich Valenz- und Leitungsband, und es gibt immer freie Zustände, um einen Elektronenfluß zu erzeugen. Wie gut nun solche Metalle den Strom leiten hängt vom Grad der Überlappung ab. Je stärker die Überlappung der Bänder, desto besser die elektrische Leitfähigkeit<sup>6</sup>.

Bei Festkörpern, die isolatorische Eigenschaften zeigen, überlappen sich die Bänder *nicht*. Sie besitzen also eine Bandlücke, und die Fermienergie (also die Energie bis zu der die Zustände mit Elektronen besetzt sind) liegt *zwischen* Valenzband und Leitungsband. Am absoluten Nullpunkt der Temperatur ist das Valenzband dann voll besetzt und das Leitungsband leer. Aber auch bei Zimmertemperatur ist die thermische Energie so klein (die thermische Energie bei  $T = 300K$  beträgt  $kT \approx 1/40eV$ ), daß es keiner nennenswerten Zahl von Elektronen gelingt, die Bandlücke von einigen bis zu  $10eV$  zu überspringen. Das Bänderschema eines Halbleiters sieht im Prinzip genau so aus wie das eines Isolators, nur ist bei Halbleitern die Bandlücke so klein (intrinsisches Silizium hat eine Bandlücke von ca.  $1.1eV$ , Germanium ca.  $0.7eV$ ), daß bei thermischen Energien im Raumtemperaturbereich oder etwas höher einige Elektronen vom Valenzband ins Leitungsband gehoben werden können und somit eine nennenswerte elektrische Leitfähigkeit zustande kommt. Für  $T \rightarrow 0$  sind Halbleiter also Isolatoren und für endliche Temperaturbereiche beginnen sie mehr und mehr den Strom zu leiten (Heißeleiter, im Gegensatz zu Metallen, deren elektrischer Widerstand sich im allgemeinen mit zunehmender Temperatur erhöht (Kaltleiter)).

## 1.2 Intrinsische Halbleiter

Wie der Name andeutet sind Halbleiter Festkörper, deren elektrische Leitfähigkeit zwischen der von Isolatoren und Metallen liegt. Der spezifische Widerstand von Halbleitermaterialien umfaßt einen Bereich von etwa  $10^{-4}\Omega m$  bis  $10^7\Omega m$ . (Im Vergleich: Metalle ca.  $10^{-7}\Omega m$ , Isolatoren  $\geq 10^{10}\Omega m$ ). Dies allein reicht aber für eine Charakterisierung nicht aus, denn es gibt z.B. Elektrolyte, deren Leitfähigkeit ebenfalls in den genannten Bereich fallen. Eine etwas genauere Definition wäre z.B. folgende:

Halbleiter sind Festkörper, die bei tiefer Temperatur isolieren, bei hohen Temperaturen jedoch meßbare *elektronische* Leitung besitzen. Diese *elektronische Leitfähigkeit* geht auf ihre wohldefinierte und durch Stromfluß unveränderliche chemische Zusammensetzung sowie auf physikalische Eingriffe von außerhalb zurück.

(aus Bergmann Schäfer[1])

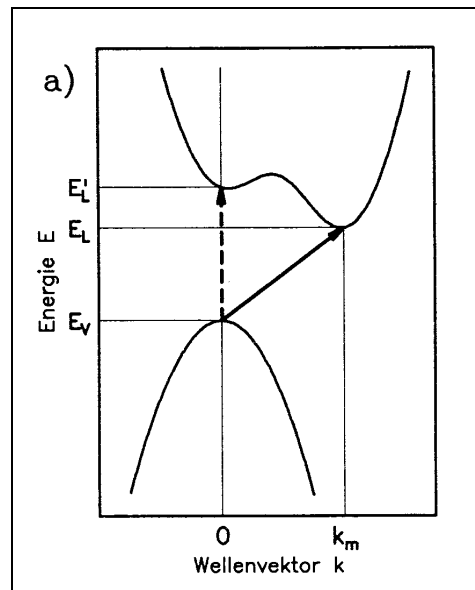
Die technisch wichtigsten Halbleitermaterialien sind Silizium und Germanium (vierte Hauptgruppe des Periodensystems der Elemente (PSE)). In letzter Zeit hat auch GaAs als eine sogenannte III–V Verbindung<sup>7</sup> an Bedeutung erlangt. Dabei ist für die Halbleitereigenschaften die „Nahordnung“ der Atomgruppen verantwortlich und weniger die „Fernordnung“, weshalb auch bei amorphen Festkörpern, Gläsern und selbst bei Flüssigkeiten Halbleitereigenschaften zu beobachten sind. Im folgenden wird fast ausschließlich von Silizium die Rede sein. Reine, bzw. *intrinsische* Halbleiter sind solche, bei denen das Material ohne, oder nur mit geringen, nicht vermeidbaren Störungen oder Verunreinigungen (engl.: impurity) vorliegt. Für technische Anwendungen sind meist nur Halbleiter mit kontrollierten Verunreinigungen, die dann *Dotierungen* genannt werden, von Bedeutung.

<sup>6</sup>Man spricht bei geringem Grad der Überlappung auch von *Halbmetallen*.

<sup>7</sup>Gallium: dritte Hauptgruppe des PSE, Arsen: fünfte Hauptgruppe des PSE

### 1.2.1 Direkte und indirekte Halbleiter

Wie zuvor erwähnt sind intrinsische Halbleiter solche, bei denen der Stoff in reiner Form vorliegt (Das ist zum Beispiel der hochreine, undotierte Wafer-Rohling). Zunächst unterscheidet man Halbleiter nach Art ihrer Bandlücke. Sie teilen sich in *direkte* und *indirekte* Halbleiter auf. Bei direkten Halbleitern liegt das Minimum des Leitungsbandes an der selben Stelle der Brioullin-Zone (bzw. beim gleichen Wellenzahlvektor  $\vec{k}$ ) wie das Maximum des Valenzbandes. Ein Elektron, welches die nötige Energie zum Überwinden der Bandlücke besitzt, kann *direkt* ins Leitungsband gelangen, da die Erhaltung des *Quasiimpulses*  $\hbar\vec{k}$  gewährleistet ist. Beim indirekten Halbleiter dagegen befindet sich das Maximum des Valenzbandes nicht an der selben Stelle des  $\vec{k}$ -Raumes wie das Minimum des Leitungsbandes. Um zum Beispiel



**Abbildung 1.3:** Schematische Bandstruktur eines indirekten Halbleiters (aus [7])

ein Photon geringstmöglicher Energie zu absorbieren (also der Energie der Bandlücke  $E_g$ ), muß am Übergang ein Phonon zur Erhaltung des Quasiimpulses beteiligt sein. Damit die Bedingungen des Energiesatzes und der Impulserhaltung erfüllt sind, müssen  $\hbar\omega_1 \pm \hbar\omega_q = E_g$  und  $\hbar\vec{k}_1 \pm \hbar\vec{q} = \hbar\vec{k}_2$  gelten. Dabei sind  $\omega_1$  die Kreisfrequenz des Photons,  $\omega_q$  die des Phonons und  $\vec{k}_1$ ,  $\vec{k}_2$ ,  $\vec{q}$  die Wellenzahlvektoren des ursprünglichen, des angehobenen Zustandes und des Phonons. Es gilt nun  $\hbar\omega_q \ll E_g$  und  $|\vec{k}_1| \ll |\vec{k}_2|$ , so daß das Photon die zum Übergang nötige Energie und das Phonon den nötigen Impuls liefert. Dadurch ist ein Übergang mit geringstmöglicher Photonenenergie beim indirekten Halbleiter weniger wahrscheinlich, als beim direkten Halbleiter, der dazu keine Beteiligung eines Phonons benötigt, da die Wellenzahlvektoren von Anfang- und Endzustand gleich sind. Erhöht man nun beim indirekten Halbleiter die eingestrahlte Photonenenergie, so können bei Erreichen des *vertikalen* Bandabstandes die Elektronen direkt, also ohne Beteiligung eines Phonons, ins Leitungsband gelangen – der Absorptionskoeffizient erhöht sich. Die Abbildung 1.3 zeigt schematisch eine Bandstruktur mit indirekter Bandlücke. Ebenfalls ist jeweils ein direkter (gestrichelt) und indirekter Bandübergang eingezeichnet. Silizium und Germanium sind Beispiele für Halbleiter mit indirekten

Bandlücken, während GaAs ein direkter Halbleiter ist.

### 1.2.2 Ladungsträgerkonzentration

Wie oben erwähnt, ist die elektrische Leitfähigkeit von Halbleitern stark temperaturabhängig. Das liegt daran, daß die Elektronen die Bandlücke zwischen dem voll besetzten Valenzband und dem leeren Leitungsband überwinden müssen. Um also die Elektronendichte im Leitungsband zu berechnen, muß die Zustandsdichte im Leitungsband mit der entsprechenden Verteilungsfunktion multipliziert und über die gesamten Energien integriert werden:

$$n = \int_{E_L}^{\infty} D_L(E) f(E, T) dE.$$

Wegen der rasch abklingenden Verteilungsfunktion  $f(E, T)$  kann die obere Grenze der Integration anstatt bis zur Oberkante des Leitungsbandes bis Unendlich gesetzt werden. Da die Löcher ebenso zum Stromtransport beitragen und deren Besetzungswahrscheinlichkeit gerade durch die fehlenden Elektronen gegeben ist, gilt:

$$p = \int_{-\infty}^{E_V} D_V(E) [1 - f(E, T)] dE.$$

Betrachtet man relativ geringe Elektronenkonzentrationen im Leitungsband, so kann man davon ausgehen, daß diese aus der Nähe des Maximum des Valenzbandes kommen. Das heißt, man kann einen parabelförmigen Verlauf des Bandes annehmen und als Zustandsdichte diejenige des freien Elektronengases annehmen<sup>8</sup>. Als zusätzliche Verbesserung wird nun aber die effektive Masse  $m^*$ <sup>9</sup> der Elektronen bzw. Löcher benutzt. Weiterhin kann die fermi'sche Verteilungsfunktion im Grenzfall  $|E_L - E_F| \gg kT$  bzw.  $|E_V - E_F| \gg kT$  durch die Boltzmann-Verteilung angenähert werden:

$$f(E, T) = \frac{1}{e^{\frac{(E-E_F)}{kT}} + 1} \stackrel{(E-E_F) \gg kT}{\approx} e^{-\frac{(E-E_F)}{kT}}.$$

Mit diesen Vereinfachungen läßt sich das Integral analytisch lösen:

$$n = \int_{E_L}^{\infty} D_L(E) f(E, T) dE = \frac{(2m^*)^{3/2}}{2\pi^2 \hbar^3} e^{E_F/kT} \int_{E_L}^{\infty} \sqrt{E - E_L} e^{-E/kT} dE.$$

Für die Elektronenkonzentration  $n$  im Leitungsband bzw. für die Löcherkonzentration  $p$  im Valenzband folgt daraus:

$$n = 2 \overbrace{\left( \frac{m_n^* kT}{2\pi \hbar^2} \right)^{3/2}}^{N_L} e^{-\frac{(E_L - E_F)}{kT}} \quad (1.1)$$

$$p = 2 \overbrace{\left( \frac{m_p^* kT}{2\pi \hbar^2} \right)^{3/2}}^{N_V} e^{\frac{(E_V - E_F)}{kT}} \quad (1.2)$$

<sup>8</sup>Zustandsdichte des freien Elektronengases siehe Anhang B.1.

<sup>9</sup>Zm Konzept der effektiven Masse siehe Anhang B.3.



Die beiden Abkürzungen  $N_L$  und  $N_V$  werden *effektive Zustandsdichten* des Leitungs- bzw. Valenzbandes genannt. Die Konzentrationen der freien Ladungsträger hängen also, bei gegebener Temperatur, entscheidend von der Lage des Fermi-Niveaus ab. Das Produkt  $n \cdot p$  dagegen ist unabhängig von der Lage des Fermi-Niveaus und ist bei fester Temperatur nur von der Bandlücke  $E_g = E_L - E_V$  abhängig:

$$np = N_L N_V e^{-E_g/kT}$$

Das Produkt der beiden Ladungsträgerkonzentrationen ist also von der Form  $n \cdot p = \text{const}$  und wird daher als *Massenwirkungsgesetz* bezeichnet. Im Falle der Eigenleitung (intrinsischer Halbleiter) gilt für die Konzentration der Ladungsträger:

$$n_i = p_i = \sqrt{np} = \sqrt{N_L N_V} e^{-E_g/2kT} \quad (1.3)$$

In Tabelle 1.1 erkennt man die abnehmende Ladungsträgerkonzentration mit zunehmender

	$E_g$ [eV]	$n_i$ [ $cm^{-3}$ ]
Germanium	0.67	$2.4 \cdot 10^{13}$
Silizium	1.1	$1.5 \cdot 10^{10}$
Galliumarsenid	1.43	$1.8 \cdot 10^6$

**Tabelle 1.1:** Bandlücken und intrinsische Ladungsträger-Konzentrationen dreier Halbleiter bei  $T = 300K$  (aus [7]).

Bandlücke. Weiterhin kann man für den Eigenleitungsfall ( $n_i = p_i$ ) die Lage des Fermi-niveaus bestimmen:

$$N_L e^{-\frac{(E_L - E_F)}{kT}} = N_V e^{\frac{(E_V - E_F)}{kT}} \iff E_F = \frac{(E_L + E_V)}{2} + \frac{kT}{2} \ln \left( \frac{N_V}{N_L} \right).$$

Wenn also die effektiven Zustandsdichten, und damit die effektiven Massen ( $N_{L,V} \propto m_{n,p}^*{}^{3/2}$ ) in Valenz- und Leitungsband gleich sind, dann ist die Lage des Fermi-niveaus unabhängig von der Temperatur und genau in der Mitte der Bandlücke<sup>10</sup>. Sind die beiden Zustandsdichten ungleich, so ist die Lage der Fermienergie durch die Bedingung  $f(E = E_F, T) = 1/2$  festgelegt. In den beiden Bildern der Abbildung 1.4 sind jeweils die Zustandsdichten für Valenz- und Leitungsband, die fermi'sche Verteilungsfunktion und das Produkt aus Zustandsdichte und Verteilungsfunktion zu sehen. Die Zustandsdichten werden hier mit  $D_V$  bzw.  $D_L$  bezeichnet. Im linken Bild ist der Fall gleicher Zustandsdichten gezeigt, das Fermi-niveau liegt in der Mitte der Bandlücke. Im rechten Bild ist  $D_V > D_L$ . Das Fermi-niveau muß wegen der Bedingung  $n_i = p_i$  in Richtung der geringeren Zustandsdichte wandern, denn die *Anzahl* der beiden Ladungsträgertypen muß im intrinsischen Fall identisch sein.

## 1.3 Dotierte Halbleiter

Die oben beschriebenen intrinsische Halbleiter kommen in der Praxis so nicht vor, denn es lassen sich bei der Herstellung von noch so reinen Kristallen gewisse Störungen nicht vermeiden.

<sup>10</sup>Zwar verengt sich die Bandlücke etwas mit zunehmender Temperatur, das Fermi-Niveau bleibt aber in der Mitte der Lücke.

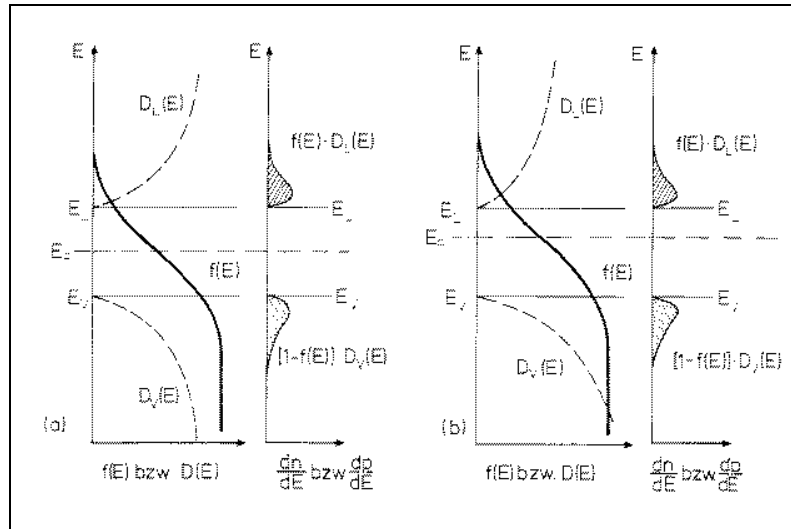
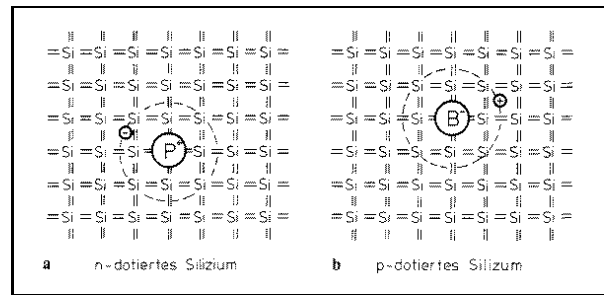


Abbildung 1.4: Lage des intrinsischen Fermi-Niveaus (aus [3])

So liegen zum Beispiel bei den reinsten GaAs-Einkristallen die Ladungsträgerkonzentrationen um die  $10^{16} \text{ cm}^{-3}$ , was deutlich von der in Tabelle 1.1 angegebenen Konzentration von  $1.8 \cdot 10^6 \text{ cm}^{-3}$  abweicht. Die dadurch erreichbare elektrische Leitfähigkeit wäre ohnehin viel zu gering für eine technische Anwendung. Um für die Ladungsträgerkonzentration bestimmte Werte zu erzielen, werden Halbleiter *dotiert*. Unter Dotierung versteht man also das kontrollierte Einbringen von *elektrisch aktiven* Fremdatomen in das Wirtsgitter.

Silizium ist vierwertig. Jedes Valenzelektron ist im Silizium-Wirtsgitter mit einem Valenzelektron des Nachbaratoms eine kovalente  $sp^3$ -Hybridbindung eingegangen. Ersetzt man nun ein Siliziumatom durch eines der fünften Hauptgruppe (z.B. Phosphor), so werden von den fünf Valenzelektronen nur vier für die Bindung benötigt. Das fünfte Elektron ist nicht abgesättigt und somit nur schwach an das Fremdatom gebunden. Es kann sich also fast frei im Gitter bewegen. Man nennt diesen Kristall dann *n-dotiert*, weil in ihm als Ladungsträger freie *Elektronen* auftreten. Ebenso kann ein reguläres Silizium Gitteratom durch ein Fremdatom der dritten Hauptgruppe (z.B. Bor) ersetzt werden. Das Boratom versucht nun, die  $sp^3$ -Bindung des Wirtsgitters mit den umgebenden Siliziumatomen einzugehen. Dazu fehlt ihm aber ein Elektron, welches es aus dem Wirtsgitter entnimmt. Dadurch fehlt im Valenzband des Siliziumkristalles ein Elektron – es ist also ein Loch entstanden. Dieser Kristall ist durch das Vorhandensein von zusätzlichen *Löchern* damit *p-dotiert*. Aufgrund der Eigenschaft, dem regulären Kristallgitter Elektronen zu *entnehmen*, nennt man die dreiwertigen Fremdatome im vierwertigen Wirtsgitter *Akzeptoren*, wogegen die fünfwertigen, die dem Leitungsband zusätzliche Elektronen *liefern*, als *Donatoren* bezeichnet werden. In beiden Fällen handelt es sich um ein wasserstoffähnliches Problem. Beim Phosphor kreist das fünfte Elektron um den einfach positiv geladenen Atomrumpf, während sich beim Bor ein Loch um den einfach negativ geladenen Rumpf bewegt. In der Abbildung 1.5 ist schematisch ein n- bzw. p-dotiertes Siliziumgitter zu sehen. Es ist zu erkennen, wie sich die beiden Fremdatome den Bindungsverhältnissen des Wirtsgitters anpassen. Ebenfalls ist das überschüssige Elektron (Loch) von Phosphor (Bor) zu sehen, dessen „Radius“ aber viel zu klein eingezeichnet ist und in Wirklichkeit ca. den zehnfachen Gitterabstand umfaßt. Um abzuschätzen, wie stark

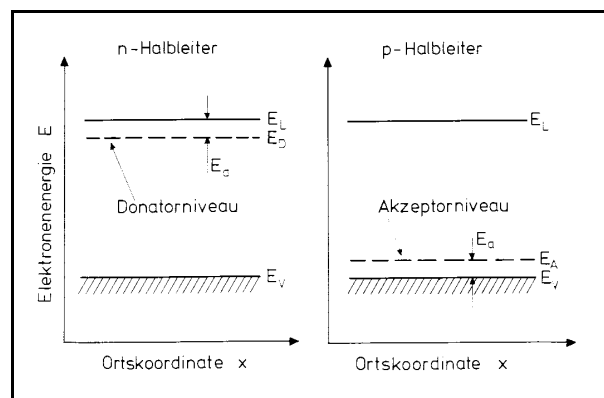


**Abbildung 1.5:** Silizium-Wirtsgitter mit Phosphor als Donator (links) und Bor als Akzeptor (aus [3])

das Elektron beim Phosphor gebunden ist (bzw. das Loch beim Bor), betrachtet man den niedrigsten Energieterm des Wasserstoff-Atomes, benutzt aber anstatt der Masse des freien Elektrons die effektive Masse eines Silizium-Elektrons, und um die Abschirmung durch das umgebende Silizium zu berücksichtigen, ersetzt man die Dielektrizitätskonstante des Vakuums durch die des Siliziums<sup>11</sup> ( $\epsilon_{Si} = 11.7$ ). Man erhält also für die „Ionisierungsenergie“ des fünfwertigen Fremdatomes:

$$13.6eV \cdot \underbrace{\frac{m_n^*}{m_e}}_{\approx 0.3} \cdot \underbrace{\frac{1}{\epsilon_{Si}^2}}_{\approx 1/137} \approx 30meV.$$

Analoge Betrachtungen gelten für die dreiwertigen Fremdatome und dem entsprechenden Loch. Ionisierungsenergie bedeutet hier die Differenz von gebundenem Donatorniveau zur



**Abbildung 1.6:** Lage der Störstellen-Niveaus bei n- und p-dotiertem Halbleiter (aus [3])

Leitungsbandkante bzw. gebundenem Akzeptorniveau zur Valenzbandkante, so daß die Energiezustände der Donatoren ca.  $30meV$  unter der der Leitungsbandunterkante und die Akzeptorniveaus ebensoviel über der Valenzbandoberkante liegen. Abbildung 1.6 zeigt die Bänder-schemen von jeweils n- und p-dotiertem Silizium inklusive der Lage der Donator- bzw. Ak-

<sup>11</sup>Die makroskopische Dielektrizitätskonstante wird dadurch gerechtfertigt, daß sich die Wellenfunktion des schwach gebundenen Elektrons wegen der Abschirmung über größenordnungsmäßig 1000 Gitteratome erstreckt.

zeptorniveaus. Da die thermische Energie bei Zimmertemperatur etwa  $25\text{meV}$  beträgt, ist einsichtig, daß sehr viele Phosphoratome bei Zimmertemperatur ionisiert sind – also freie Ladungsträger im Leitungsband des Kristalles existieren. In der Tabelle 1.2 kann man erkennen,

	Donatoren			Akzeptoren		
	P	As	Sb	B	Al	Ga
Silizium	45	49	39	45	57	65
Germanium	12	12.7	9.6	10.4	10.2	10.8

**Tabelle 1.2:** Ionisierungsenergien einiger Donatoren und Akzeptoren in Silizium und Germanium in [ $\text{meV}$ ] (aus [7])

daß sich die Ionisierungsenergien je nach Element ein wenig unterscheiden, dennoch nähert das „Wasserstoffmodell“ die wahren Gegebenheiten recht gut<sup>12</sup>.

### 1.3.1 Ladungsträgerkonzentration

Befinden sich in einem Halbleiter Donatoren *und* Akzeptoren in gleicher Konzentration, so heben sich deren Wirkungen gegenseitig auf. Dies wird auch als *Kompensation* bezeichnet. Die überschüssigen Elektronen der Donatoren besetzen dann gerade die Löcher der Akzeptoren, so daß die Dotierungen nicht zur Erhöhung der freien Ladungsträger beitragen. Der für die Mikroelektronik interessantere Fall ist derjenige, bei dem innerhalb eines bestimmten Gebietes die eine oder andere Dotierungsart dominiert<sup>13</sup>. Es werden hier also ausschließlich Halbleiter betrachtet, bei denen eine Dotierungsart um Größenordnungen überwiegt, so daß etwaige andere Verunreinigungen vollständig kompensiert sind.

Die Donatoren sind entweder mit einem Elektron besetzt oder sie sind nicht besetzt, das heißt sie sind einfach positiv ionisiert. Analog dazu sind die Akzeptoren entweder mit einem Loch besetzt, sie sind dann elektrisch neutral, oder sie haben ein Elektron aus dem Valenzband an sich gezogen, wodurch sie dann einfach negativ geladen sind.

Donatoren:

$$\underbrace{n_D}_{\text{Donatorkonzentration}} = \underbrace{n_D^0}_{\text{neutrale Donatoren}} + \underbrace{n_D^+}_{\text{ionisierte Donatoren}}$$

Akzeptoren:

$$\underbrace{n_A}_{\text{Akzeptorkonzentration}} = \underbrace{n_A^0}_{\text{neutrale Akzeptoren}} + \underbrace{n_A^-}_{\text{ionisierte Akzeptoren}}$$

Die Besetzung der Donatoren mit Elektronen ( $e^-$ ) bzw. der Akzeptoren mit Löchern<sup>14</sup> ( $e^+$ ) wird wieder durch die Fermi-Dirac Statistik geregelt:

$$n_D^0 = n_D \left[ \frac{1}{e^{\frac{(E_D - E_F)}{kT}} + 1} \right], \quad n_A^0 = n_A \left[ \frac{1}{e^{\frac{(E_F - E_A)}{kT}} + 1} \right].$$

<sup>12</sup>Man kann auch mit anderswertigen Atomen dotieren, wobei sich dann möglicherweise mehrere Niveaus ausbilden. Die technisch am wichtigsten sind aber die drei- und fünfwertigen.

<sup>13</sup>Will man z.B. hochohmige Widerstände erzeugen, so könnte man die nicht vermeidbaren Verunreinigungen durch entgegengesetzte Dotierung kompensieren und so die Leitfähigkeit kontrolliert reduzieren.

<sup>14</sup>Die *besetzten* Störstellen sind elektrisch neutral!

$E_D$  und  $E_A$  bezeichnen das Donator- bzw. Akzeptorniveau. Zu den „freien“ Ladungsträgerdichten  $n$  bzw.  $p$  tragen sowohl die ionisierten Dotieratome als auch die intrinsischen, vom Valenzband ins Leitungsband gewechselten, Elektronen und die dadurch im Valenzband verbleibenden Löcher bei.

Das Fermi-Niveau wird durch die Dotierung beeinflusst.

In Abschnitt 1.2.2 wurde die Ladungsträgerkonzentration berechnet. Dabei wurde keine Annahme für intrinsische Halbleiter gemacht. Daher gelten die Gleichungen (1.1, 1.2) und das Massenwirkungsgesetz (Gl. (1.3)) auch bei dotierten Halbleitern. Das muß dann Konsequenzen für das Fermi-Niveau haben, denn es stehen ja jetzt durch die Dotierung offensichtlich mehr Ladungsträger zur Verfügung. Dazu kommt noch eine Gleichung, die die Ladungsneutralität innerhalb des Halbleiters ausdrückt.

Die Gleichungen für einen n-Halbleiter:

$$\rho = q(p + n_D^+ - n - n_A^-) = 0 \quad (1.4)$$

$$n_D = n_D^0 + n_D^+ \quad (1.5)$$

$$n = N_L e^{\frac{-(E_L - E_F)}{kT}} \quad (1.6)$$

$$\frac{n_D^0}{n_D} = \frac{1}{e^{\frac{(E_D - E_F)}{kT}} + 1} \quad (1.7)$$

Setzt man in Gleichung 1.4 für  $p = n_i^2/n$ , so erhält man  $n^2 - (n_D^+ - n_A^-)n - n_i^2 = 0$ . Aufgelöst nach  $n$  ergibt diese Gleichung:

$$n_n = \frac{n_D^+ - n_A^- + \sqrt{(n_D^+ - n_A^-)^2 + 4n_i^2}}{2},$$

wobei  $n_n$  die Elektronenkonzentration im n-Gebiet bezeichnet. Die Elektronen in einem n-Halbleiter bezeichnet man als *Majoritäten*. Im Gegensatz dazu steht  $p_n$  für die Löcher im n-Halbleiter, deren Dichte dort sehr klein ist und deshalb *Minoritäten* genannt werden (das gilt natürlich umgekehrt für  $p_p$  – die Löcher im p-Halbleiter –, die dort die Majoritäten sind und für die Elektronen  $n_p$  die dann die Minoritäten sind). Bei der Betrachtung von dotierten Halbleitern kann man stets davon ausgehen, daß die intrinsische Ladungsträgerkonzentration sehr viel kleiner als die Dichte der ionisierten Störstellen ist ( $n_i \ll |n_D^+ - n_A^-|$ ). Daher vereinfacht sich der Ausdruck zu  $n_n = n_D^+ - n_A^-$  für die Majoritäten und durch Anwendung des Massenwirkungsgesetzes  $n_n p_n = n_i^2$ , gilt für die Minoritäten  $p_n = \frac{n_i^2}{n_D^+ - n_A^-}$ . Außerdem sollen hier Halbleiter untersucht werden, bei denen eine Dotierungsart innerhalb eines Gebietes überwiegt, das heißt es kann davon ausgegangen werden, daß im n-Halbleiter  $n_D^+ \gg n_A^-$  gilt. Dies liefert mit analogen Betrachtungen für den p-Halbleiter:

	Majoritäten	Minoritäten
n-Halbleiter	$n_n \approx n_D^+$	$p_n \approx \frac{n_i^2}{n_D^+}$
p-Halbleiter	$p_p \approx n_A^-$	$n_p \approx \frac{n_i^2}{n_A^-}$

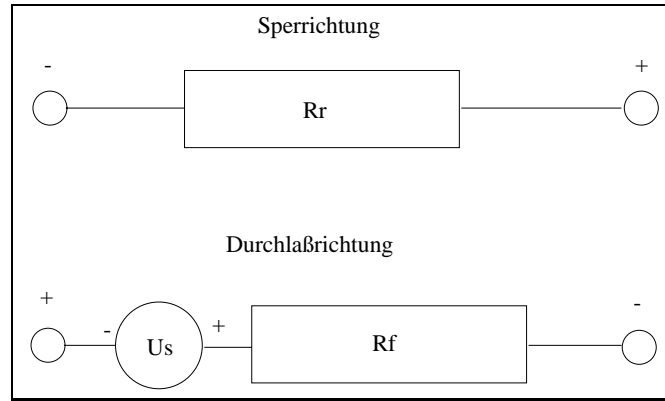
An den Formeln für die Minoritäten kann man sehen, daß in einem n-Halbleiter zum Beispiel so gut wie keine Löcher existieren. Setzt man für  $n_i$  die intrinsische Ladungsträgerkonzentration für Silizium ( $1.5 \cdot 10^{10} \text{ cm}^{-3}$ ) aus Tabelle 1.1 ein und geht von einer relativ schwachen n-Dotierung von  $10^{16} \text{ cm}^{-3}$  aus, so sind die Minoritäten nur mit einer Dichte von  $2.25 \cdot 10^4 \text{ cm}^{-3}$  vertreten.

## Kapitel 2

# Numerische Methoden

Um das Verhalten von elektronischen Bauelementen in Schaltkreisen untersuchen zu können, werden diese durch Modelle beschrieben. In einem Schaltkreis soll zum Beispiel die gleichrichtende Wirkung einer Diode beschrieben werden. Im Sperrbetrieb kann die Diode dann als hoher Ohm'scher Widerstand betrachtet werden. Wird die Diode umgepolt, so muß ihr Widerstand aber wesentlich kleiner sein. Da in Wirklichkeit der Durchlaßwiderstand einer Diode aber von der angelegten Vorwärtsspannung abhängt, muß der gewählte Ersatzwiderstand für die Beschreibung bei entsprechenden Genauigkeitsansprüchen für jeden Spannungsbereich neu gewählt werden. Das Modell kann beliebig verfeinert werden. Der Schwellenspannung in Vorwärtsrichtung kann Rechnung getragen werden, indem man dem Widerstand eine Spannungsquelle mit entgegengesetztem Vorzeichen wie die angelegte Durchlaßspannung vorschaltet. Handelt es sich um eine Siliziumdiode, so wird diese Spannung ca.  $0.7V$  betragen, bei einer Germaniumdiode ca.  $0.3V$ . Die Parameter, wie etwa der Rückwärts-Sättigungsstrom oder die Schwellenspannung müssen entweder anderweitig berechnet, oder durch Messungen erhalten werden, so daß das jeweilige Modell an die wahren Bauelemente gefittet werden kann. Die Abbildung 2.1 zeigt zwei einfache Ersatzschaltbilder für eine Diode in Sperr- und Durchlaßrichtung. Der Widerstand  $R_r$ , der den Strom, der durch die Diode in Sperrichtung fließt, modellieren soll, ist um Größenordnungen größer als  $R_f$ , der außerdem für verschiedene Vorwärtsspannungen verschiedene Werte annimmt. Desweiteren wird mit der inneren Spannungsquelle die Schwellenspannung der Diode in Vorwärtsrichtung berücksichtigt. Das Verhalten der Diode als Schaltungselement muß also vorher schon bekannt sein, um es in diesem Modell zu beschreiben. Man kann daher nicht erwarten, neue Erkenntnisse über das Bauelement selbst zu erfahren, denn diese müßten ja schon vorher im Modell enthalten sein. Diese Art von Beschreibung von elektrischen Bauelementen ist vielmehr geeignet, das Zusammenspiel mehrerer Bauelemente in einer elektronischen Schaltung zu untersuchen. Will man zum Beispiel noch Aussagen über das zeitliche Verhalten der Diode in einer Schaltung wissen, so müssen weiterhin Zeitkonstanten in Form von Kapazitäten usw. eingeführt werden. Da man es in elektronischen Schaltkreisen im allgemeinen mit sehr vielen Bauelementen zu tun hat, ist die Rechenzeit, mit der das Modell berechnet wird, von Bedeutung. Man nimmt daher gewisse Einschränkungen in der Exaktheit der Kennlinien in Kauf und rechnet mit einfacheren, schnelleren Modellen, die auch je nach Anwendung variieren.

Dem gegenüber steht das Konzept der *Bauelementen-Simulation*. Bei dieser Art der Simulation elektronischer Bauelemente gehen physikalisch grundlegendere Annahmen ein. Es



**Abbildung 2.1:** Einfaches Ersatzschaltbild einer Diode in Sperrichtung (oben) und Durchlaßrichtung

sollen dann zum Beispiel Größen, wie der Rückwärts-Sättigungsstrom  $I_r$ , der im oben beschriebenen Modell in Form des Sperrwiderstandes  $R_r$  eingeht, als Ergebnis herauskommen. Der Bauelement-Simulator wird zum Beispiel mit der Geometrie der p- und n-Zone und den Störstellenkonzentrationen „gefüttert“. Simuliert man nun von außen angelegte Spannungen, so soll der Simulator beispielsweise die Potentialverteilung  $\Psi(x, y, z)$  in diesem inhomogenen Halbleiter berechnen. Daraus kann auf die Ladungsträgerverteilung innerhalb des Halbleiters geschlossen werden, und damit zum Beispiel auf die Dicke der Raumladungszone. Beim Bauelement-Simulator sollen also im wesentlichen Daten des Herstellungsprozesses eingehen, so daß man zum Beispiel untersuchen kann, wie die Schwellenspannung einer Diode von der Höhe der Störstellenkonzentration abhängt. Weiterhin kann untersucht werden, wie sich kompliziertere Dotierprofile zum Beispiel in der pn-Grenzschicht auf das elektrische Feld in der Verarmungszone auswirken. Wie also der Name andeutet legt der Device-Simulator das Gewicht mehr auf das Bauelement selbst, als auf das Zusammenwirken verschiedener Schaltungsteile miteinander. Er könnte dazu benutzt werden, um etwa die Parameter eines vorhandenen Modelles zu berechnen. Im folgenden Abschnitt sollen nun die Grundgleichungen gezeigt werden, mit denen die elektronischen Eigenschaften von Halbleiterbauelemente berechnet werden.

## 2.1 Grundgleichungen

Die grundlegendste Gleichung der Elektrostatik ist die Poisson-Gleichung, eine Differentialgleichung zweiter Ordnung, die zu einer gegebenen Ladungsverteilung die dazugehörige Potentialverteilung liefert.

Poisson-Gleichung:

$$\operatorname{div}(\epsilon \nabla \Psi) = -\rho = -\epsilon (\operatorname{div} \vec{E}), \quad (2.1)$$

dabei bedeuten:

$\operatorname{div}$  den einfachen Ableitungsoperator für ein Vektorfeld  $\vec{V}(\vec{r})$  ( $\operatorname{div} \vec{V}(\vec{r}) = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$ ),



$\epsilon$  die Dielektrizitätskonstante im Vakuum,

$\nabla$  den Gradienten eines skalaren Feldes  $\phi(\vec{r})$  ( $\nabla\phi(\vec{r}) = \left(\frac{\partial\phi(\vec{r})}{\partial x}, \frac{\partial\phi(\vec{r})}{\partial y}, \frac{\partial\phi(\vec{r})}{\partial z}\right)$ ),

$\Psi$  das elektrostatische Potential, und

$\rho(\vec{r})$  die Ladungsdichteverteilung im Halbleiter.

Für die elektrische Feldstärke  $\vec{E}$  gilt:

$$\vec{E} = -\nabla\Psi. \quad (2.2)$$

Als zweite grundlegende Gleichung kommt die Kontinuitätsgleichung, hier getrennt dargestellt in die für Elektronen und Löcher, hinzu.

Kontinuitätsgleichung:

$$\frac{\partial n}{\partial t} = \frac{1}{q} \operatorname{div} \vec{J}_n + G_n - R_n, \quad (2.3)$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \operatorname{div} \vec{J}_p + G_p - R_p, \quad (2.4)$$

mit den Bezeichnungen

$n, p$  für die Elektronen bzw. Löcherkonzentration,

$q$  für die Elementarladung,

$\vec{J}_n, \vec{J}_p$  für die Stromdichten der Elektronen und Löcher,

$G_n, G_p$  für die Generationsrate von Elektronen und Löcher, und

$R_n, R_p$  für die Rekombinationsraten.

Diese Gleichungen sind das Herzstück des Simulators. Damit alleine läßt sich natürlich nicht rechnen. Es müssen weitere Gleichungen folgen, die die Bedingungen näher spezifizieren. Zum Beispiel muß festgelegt werden, wie die Ladungsdichte  $\rho(\vec{r})$  aussieht. Ebenso müssen Modelle gefunden werden, die die Rekombinationsraten, Generationsraten usw. bestimmen. Weil diese Gleichungen in diesem Sinne keine physikalischen Grundgleichungen mehr darstellen, sondern mehr oder weniger aus empirischen Daten und Meßwerten hervorgehen, werden sie hier als *sekundäre* Gleichungen bezeichnet.

## 2.2 Sekundäre Gleichungen

Im Gegensatz zu den obigen Differentialgleichungen, die analytisch und exakt sind, folgen jetzt solche, bei denen unter gewissen physikalischen Bedingungen Näherungen gemacht werden, einerseits um überhaupt auf ein numerisches Ergebnis zu kommen, andererseits um die Berechnungen effizienter und schneller zu machen. Weitere Gleichungen, die die Physik des Halbleiters beschreiben, betreffen die Besetzungswahrscheinlichkeit, die effektiven Zustandsdichten und die intrinsische Ladungsträgerkonzentration. Die Besetzung der Energiezustände mit Elektronen im Halbleiter wird durch die Fermi-Dirac Statistik geregelt. Die Fermi-Dirac Funktion geht im Grenzfalle  $E - E_F \gg kT_L$  in die einfacher zu berechnende Boltzmann-Statistik über. Da die Näherung für die meisten Fälle in der Halbleiterelektronik bei Raumtemperatur gegeben ist, wird die Besetzung der Energiezustände meist mit der

Boltzmann-Verteilung berechnet. Das gilt allerdings nicht mehr für hochdotierte sogenannte *degenerierte* Halbleiter, wo dann explizit die Fermi-Dirac Verteilung benutzt werden muß. Die Näherung bricht dann zusammen, weil sich das Fermi-Niveau bei hoher Dotierung bis in das Leitungsband schiebt und dann nicht mehr nur der „Ausläufer“ der Fermi-Dirac Verteilung mit Elektronen besetzt ist. Es macht sich dann die Natur der Elektronen als Fermionen bemerkbar, wo dann zwei identische Teilchen nicht einen gemeinsamen Zustand einnehmen dürfen (Pauli-Prinzip).

Fermi-Dirac bzw. Boltzmann-Verteilung:

$$f(E, T) = \frac{1}{1 + e^{\frac{(E-E_F)}{kT_L}}} \stackrel{E-E_F \gg kT_L}{\approx} e^{-\frac{(E_F-E)}{kT_L}} \quad (2.5)$$

Dabei bedeuten:

$E$  ein Energiezustand, der mit einem Elektron besetzt werden kann,

$E_F$  die Fermi-Energie (also das chemische Potential  $\mu$  für  $T \rightarrow 0$ ),

$k$  die Boltzmann-Konstante und

$T_L$  die Temperatur des Festkörperlattices (engl.: lattice-temperature).

Ladungsträgerkonzentrationen:

$$n = N_C e^{\frac{(E_F-E_C)}{kT_L}}, \quad (2.6)$$

$$p = N_V e^{\frac{(E_V-E_F)}{kT_L}}. \quad (2.7)$$

$n, p$  bedeuten hier die Konzentrationen der Elektronen und Löcher,

$E_C$  die Energie des Leitungsbandkantenminimums (engl.: conduction band),

$E_V$  die Maximalenergie der obersten Valenzbandkante.

Effektive Zustandsdichten:

$$N_C = 2 \left[ \frac{2\pi m_e^* kT_L}{h^2} \right]^{3/2} \quad (2.8)$$

$$N_V = 2 \left[ \frac{2\pi m_h^* kT_L}{h^2} \right]^{3/2} \quad (2.9)$$

$N_C$  bzw.  $N_V$  in (2.6) und (2.7) werden als effektive Zustandsdichten bezeichnet,

$m_e^*$  und  $m_h^*$  als effektive Massen<sup>1</sup> der Leitungsbandelektronen bzw. Valenzbandlöcher.

$h$  ist das Planck'sche Wirkungsquantum.

Die effektive Zustandsdichte ist eine Näherung über einen parabolischen Verlauf der Bandstruktur, also der  $E(\vec{k})$ -Kurve. Der Zusammenhang  $E \propto k^2$  gilt für das freie Elektronengas. Da aber hauptsächlich Elektronen im Minimum des Leitungsbandes und Löcher im Maximum des Valenzbandes betrachtet werden, lassen sich die Bänder an diesen Stellen recht gut

<sup>1</sup>Zm Konzept der effektiven Masse siehe Anhang B.3.

durch einen parabolischen Verlauf nähern. Die effektiven Massen  $m_{e,h}^*$  der Elektronen und Löcher (engl.: holes) werden durch Mittelung über die verschiedenen Bandkrümmungen in verschiedene Richtungen der Brioullin-Zone im Halbleiterkristall erhalten ( $\frac{1}{m_{e,h}^*} \propto \frac{\partial^2 E(\vec{k})}{\partial k^2}$ ).

Eine weitere wichtige Beziehung in der Halbleiterphysik ist die der intrinsischen Ladungsträgerkonzentrationen. Multipliziert man (2.6) mit (2.7), so findet man, daß das Produkt  $n \cdot p$  der Ladungsträgerdichten bei gegebener Temperatur eine Konstante ist, die nur von der Bandlücke des jeweiligen Halbleiters abhängt.

Intrinsische Ladungsträgerkonzentration:

$$n \cdot p = n_i^2 = \text{const} \quad (2.10)$$

$$n_i = \sqrt{N_C N_V} e^{\frac{-E_g}{2kT_L}} \quad (2.11)$$

$n_i$  bezeichnet die intrinsische Ladungsträgerkonzentration, also die im undotierten Halbleiter.  $E_g$  ist die Bandlücke des Kristalls (engl.: bandgap).

Für den intrinsischen Halbleiter gilt  $n = p$ , da alle Elektronen im Leitungsband aus dem Valenzband des Halbleiters kommen und damit dort ein Loch hinterlassen. Setzt man also (2.6) mit (2.7) gleich, so ergibt sich sofort die Lage des intrinsischen Fermi-Niveaus:

$$E_F = E_i = -q\Psi_i = \frac{(E_C + E_V)}{2} + \frac{kT_L}{2} \ln \left( \frac{N_V}{N_C} \right) \quad (2.12)$$

Hier bedeuten

$E_i$  das intrinsische Fermi-Niveau,

$q$  der Betrag der Elementarladung, und

$\Psi_i$  das intrinsische Potential.

Die Elektronen- und Löcherkonzentrationen können durch die intrinsische Ladungsträgerkonzentration ausgedrückt werden:

$$n = n_i e^{\frac{q(\Psi - \Phi)}{kT_L}} \quad (2.13)$$

bzw.

$$p = n_i e^{\frac{q(\Phi - \Psi)}{kT_L}}, \quad (2.14)$$

dabei sind

$\Psi$  das intrinsische Potential und

$\Phi$  das Potential, das dem Fermi-Niveau entspricht ( $\Phi = -\frac{E_F}{q}$ ).

In Gleichung (2.11) stehen als Materialparameter die Bandlücke  $E_g$  und die effektiven Zustandsdichten des Valenz- und Leitungsbandes  $N_V$  und  $N_C$ , um die intrinsische Ladungsträgerkonzentration zu erhalten. Um die Temperaturabhängigkeit dieser Größen, die in Gleichung (2.11) als konstant angenommen wurden, zu berücksichtigen, werden sie mit folgenden Formeln modelliert:

$$E_g(T_L) = E_g(300) + \alpha \left[ \frac{300^2}{300 + \beta} - \frac{T_L^2}{T_L + \beta} \right] \quad (2.15)$$

$$N_C(T_L) = \left(\frac{T_L}{300}\right)^{3/2} N_C(300) \quad (2.16)$$

$$N_V(T_L) = \left(\frac{T_L}{300}\right)^{3/2} N_V(300) \quad (2.17)$$

Die Parameter  $N_C(300)$ ,  $N_V(300)$ ,  $E_g(300)$ ,  $\alpha$  und  $\beta$  können je nach verwendetem Material angegeben werden.

Weiterhin wird die Bandlückenverengung durch *hohe* Dotierung dadurch berücksichtigt, indem die intrinsische Ladungsträgerkonzentration aus Gleichung (2.11) durch eine effektive Ladungsträgerdichte ersetzt wird:

$$n_{ie} = n_i \exp \left\{ \frac{qa_1}{kT} \left[ \ln \frac{N(x,y)}{a_2} + \left( \left( \ln \frac{N(x,y)}{a_2} \right)^2 + a_3 \right)^{1/2} \right] \right\} \quad (2.18)$$

Die Parameter  $a_1$ ,  $a_2$  und  $a_3$  können wieder je nach Material spezifiziert werden.

Hier wird deutlich, warum es nützlich ist, die Dichte der Ladungsträger durch die der intrinsischen Ladungsträgerkonzentration auszudrücken, denn dadurch wird die Erhöhung der Ladungsträger durch die Bandlückenverengung bei hoher Dotierung automatisch berücksichtigt. Da der Effekt relativ klein ist und auch erst bei hoher Dotierung wirksam wird, wird das Modell nur bei Bedarf zugeschaltet. Standardmäßig wird also der Einfachheit halber bzw. mit Rücksicht auf die Rechenzeit, die Bandlückenverengung nicht modelliert.

Die Gleichungen (2.6) und (2.7) für die Ladungsträgerdichten gelten im *thermodynamischen Gleichgewicht*. Der Simulator soll aber gerade die elektrischen Eigenschaften von Halbleiterbauelementen, die einem externen elektrischen Feld unterliegen, berechnen. Diese befinden sich dann nicht mehr im thermodynamischen Gleichgewicht.

Um auch für diese Bedingungen ähnliche Ausdrücke zu erhalten, werden die sogenannten *Quasi-Fermi Levels* eingeführt:

$$n = n_{ie} e^{\frac{q(\Psi - \Phi_n)}{kT_L}} \iff \Phi_n = \Psi - \frac{kT_L}{q} \ln \frac{n}{n_{ie}} \quad (2.19)$$

bzw.

$$p = n_{ie} e^{\frac{q(\Phi_p - \Psi)}{kT_L}} \iff \Phi_p = \Psi + \frac{kT_L}{q} \ln \frac{p}{n_{ie}} \quad (2.20)$$

Dabei bezeichnen  $\Phi_n$ ,  $\Phi_p$  die Quasi-Fermi Levels für die Elektronen bzw. Löcher.

Die Stromdichten werden dann durch die Quasi-Fermi Levels ausgedrückt:

$$\vec{J}_n = -q\mu_n n \nabla \Phi_n \quad (2.21)$$

und

$$\vec{J}_p = -q\mu_p p \nabla \Phi_p \quad (2.22)$$

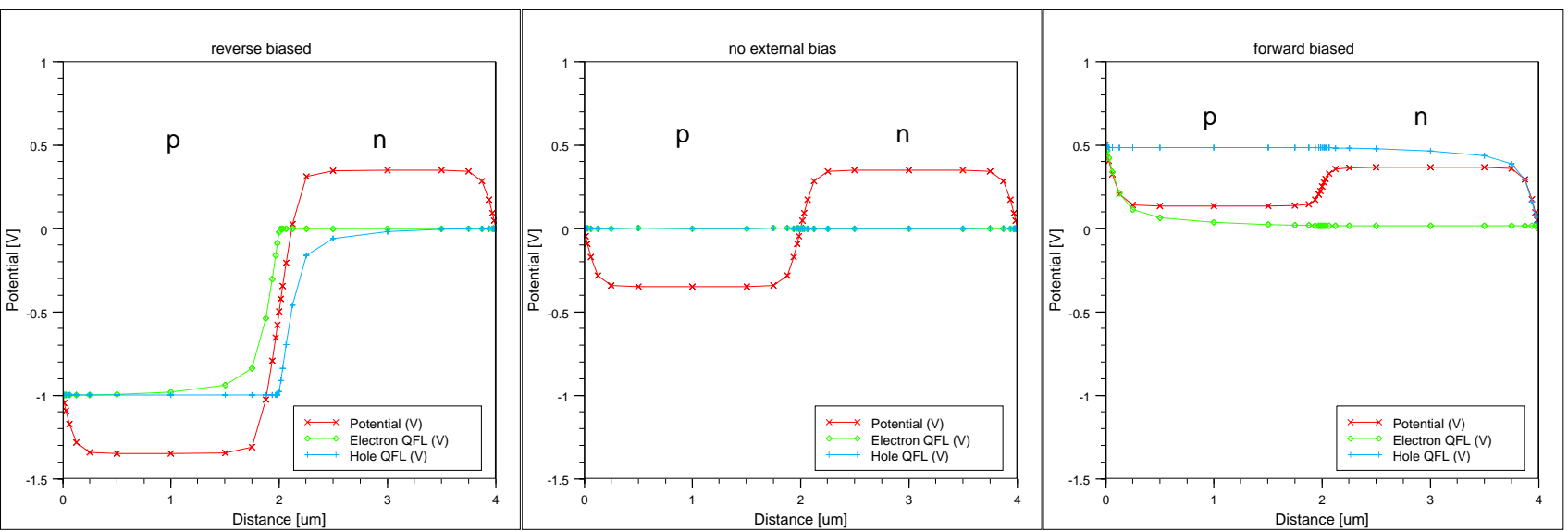
$\mu_n$  und  $\mu_p$  sind die Beweglichkeiten<sup>2</sup> der Elektronen und Löcher.

In der Poisson Gleichung (2.1) wird für das Potential *immer* das intrinsische Potential  $\Psi$  benutzt. Im thermodynamischen Gleichgewicht stellt sich das Fermi-Potential gemäß den Gleichungen (2.13 und 2.14) zu  $\Phi = \Psi - \frac{kT}{q} \ln \frac{n}{n_i}$  ein. Hierbei entsprechen dem intrinsischem Potential das intrinsische Fermi-Niveau  $-q\Psi = E_i$  und dem Potential das Fermi-Niveau  $-q\Phi = E_F$ , welches sich zum Beispiel durch Dotierung und der damit verbundenen Erhöhung der verfügbaren Ladungsträger einstellt. Im Gleichgewicht liegen die Fermi-Energien im kompletten Kristall immer auf gleicher Höhe. Das heißt, es wird auch nicht zwischen einer Fermi-Energie für Löcher und Elektronen unterschieden. Das bewirkt, durch die unterschiedlichen Abstände der Fermi-Energien zu den Bandkanten in n- und p-Halbleitern, die Bandverbiegung an der Kontaktzone. Unterwirft man nun den Halbleiter äußeren elektrischen Feldern, so gilt Gleichung (2.10) nicht mehr. Die Konzentration der Ladungsträger hat sich durch Einwirkung einer äußeren Energie geändert. Das betrifft vor allem die Minoritäten, deren Konzentration sich merklich verändern läßt, während sich an den Konzentrationen der Majoritäten prozentual gesehen nicht viel ändert. Weil aber die Situation nicht allzu weit von der Gleichgewichtslage entfernt ist, beschreibt man die Änderung der Ladungsträgerkonzentration durch das Konzept der Quasi-Fermi Levels. Diese sind dann für die Elektronen und Löcher nicht mehr gleich. Die Quasi-Fermi Levels stellen sich nun so ein, daß gerade die neue Ladungsträgerkonzentration berücksichtigt wird. Bildet man wie in Gleichung (2.10) das Produkt aus der Konzentration der Elektronen und der der Löcher mit den Dichten aus den Gleichungen (2.19 und 2.20), so erhält man  $n \cdot p = n_{ie}^2 \exp\{\frac{q(\Phi_p - \Phi_n)}{kT}\}$ . Man sieht daran, daß für  $\Phi_p > \Phi_n$  das Produkt  $n \cdot p > n_{ie}^2$  wird und umgekehrt für  $\Phi_p < \Phi_n$  das Produkt  $n \cdot p < n_{ie}^2$ . Im ersten Fall ist der pn-Übergang in Vorwärtsrichtung gepolt, während im zweiten eine Rückwärtsspannung anliegt. Die Quasi-Fermi Levels entsprechen also den von außen angelegten Spannungen an p- und n-Zone.

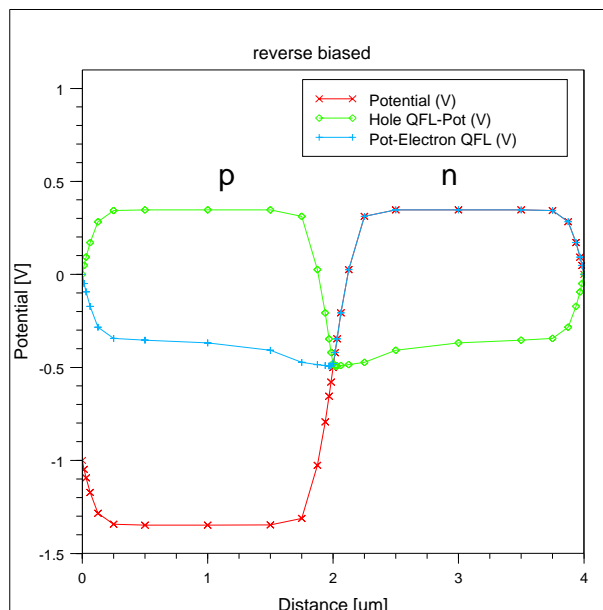
Der Simulator setzt also die externen Potentiale den Quasi-Fermi Levels gleich und berechnet sich daraus die Ladungsträgerkonzentration gemäß den Gleichungen (2.19 und 2.20). Diese fließt dann in die Ladungsdichte  $\rho(\vec{r})$  der Poissongleichung (2.1) ein, mit der dann das intrinsische Potential  $\Psi$  im Halbleiterkristall berechnet wird. Aufgrund der Materialeigenschaften werden das Valenz- und Leitungsband generiert.

Um die Potentialverhältnisse zu verdeutlichen, zeigt Abbildung 2.2 das intrinsische Potential (im Bild als *Potential* bezeichnet (rote Kurve)) und die Quasi-Fermi Levels der Elektronen (Electron QFL (grün)) und Löcher (Hole QFL (blau)) einer symmetrisch abrupt dotierten Diode mit einer Phosphor-Dotierung von  $1 \cdot 10^{16}/cm^3$  im n-Gebiet ( $2 \leq x \leq 4$ ) und einer gleich großen Konzentration von Bor im p-Gebiet ( $0 \leq x < 2$ ). Die Dotiergebiete stoßen bei  $x = 2\mu m$  aufeinander. Jeweils links in den Bildern befindet sich also das p-Gebiet. Es soll zunächst das linke Bild betrachtet werden, bei dem am p-Gebiet eine negative Spannung von  $-1V$  anliegt und der Kontakt an der Kathode geerdet ist ( $0V$ ). Nähert man sich nun von links kommend der Verarmungszone ( $x \leq 2\mu m$ ), so sieht man am Fermi-Potential der Löcher, die dort die Majoritäten sind, daß sich deren Konzentration aufgrund der äußeren Spannung in diesem Gebiet nicht ändert. Im Gegensatz dazu steigt aber der Quasi-Fermi Level für die Elektronen kurz vor der Verarmungszone an, so daß sich die Differenz  $\Psi - \Phi_n$  zwischen dem intrinsischen Potential und dem Quasi-Fermi Level der Elektronen verkleinert, was nach Gleichung (2.19) eine Verringerung der Ladungsträger bedeutet. Die Verarmungszone wird

<sup>2</sup>Zur Modellierung der Beweglichkeiten siehe Abschnitt 2.3



**Abbildung 2.2:** Potentiale einer symmetrisch abrupt dotierten Diode bei verschiedenen äußeren Spannungen (Erklärung siehe Text)



**Abbildung 2.3:** Diode in Sperrichtung gepolt (Erklärung siehe Text)

also größer. Analoge Betrachtungen gelten für die Löcher im n-Gebiet. Die Abbildung 2.3 zeigt nochmals dieselbe Diode in Sperrichtung gepolt, hier sind aber anstatt der Quasi-Fermi Levels die Differenzen  $\Psi - \Phi_n$  und  $\Phi_p - \Psi$  (im Bild als *Pot-Electron QFL* (blau) und *Hole QFL-Pot* (grün) bezeichnet) eingezeichnet. Diese Differenzen sind nach Gleichungen (2.19 und 2.20) proportional dem Logarithmus der entsprechenden Ladungsträgerkonzentrationen, so daß man hier direkt die Absenkung der Minoritätendichte nahe der Verarmungszone ablesen kann. Im mittleren Bild der Abbildung 2.2 ist der Fall des thermischen Gleichgewichtes dargestellt. Es liegen keine äußere Spannungen an. Wie oben erwähnt fallen nun die Quasi-Fermi Levels zusammen. Im Falle einer Durchlaßspannung, wie im rechten Bild gezeigt, gibt es wieder unterschiedliche Fermi-Niveaus für Elektronen und Löcher. Wieder werden nur die Minoritäten wesentlich von den externen Spannungen beeinflusst. Im p-Gebiet zum Beispiel ist der Abstand von Quasi-Fermi Level der Löcher zum intrinsischen Potential in allen drei Bildern gleich, so daß sich an deren Konzentrationen bei verschiedenen äußeren Spannungen nichts ändert. Das gleiche gilt für die Elektronen im n-Gebiet.

Die Potentialdifferenzen erhält man am einfachsten, indem man zuerst das thermische Gleichgewicht betrachtet (mittleres Bild). Das intrinsische Potential liegt nach Gleichung (2.12) in der Mitte zwischen Leitungs- und Valenzband, sofern man  $T \rightarrow 0$  betrachtet. Bei endlichen Temperaturen hängt die Lage des intrinsischen Fermi-Niveaus von den effektiven Zustandsdichten ab. Für Silizium gilt  $\ln\left(\frac{N_V}{N_C}\right) \approx 0.99$ , so daß auch hier bei endlichen Temperaturen die intrinsische Fermi-Energie in der Mitte der Bandlücke liegt. Der Abstand des Fermi-Potentials der n-Zone, welches sich durch die Dotierung einstellt, zum intrinsischen Potential ergibt sich aus den Gleichungen (2.10, 2.6) und der Annahme, daß alle Donatoren ionisiert sind<sup>3</sup> und die Störstellenleitung überwiegt, zu  $E_F - E_i = kT \ln\left(\frac{N_D}{n_i}\right)$ . Setzt man die Do-

<sup>3</sup>Das ist bei Raumtemperatur eine gute Annahme.

tierung der Diode ein, so erhält man  $E_F - E_i \approx 0.35eV$ . Analoge Betrachtungen führen im p-Gebiet zu  $E_i - E_F \approx 0.35eV$ . Bei gegebener Temperatur ändert sich die Differenz zwischen dem intrinsischen Fermi-Niveau und dem Fermi-Niveau, welches durch die Dotierung gegeben ist, nicht. Insbesondere ändert sie sich nicht, wenn der Halbleiter äußeren elektrischen Feldern unterworfen ist (sofern sich dadurch seine Gittertemperatur nicht wesentlich erhöht). Es ändert sich aber die Elektronenkonzentration, so daß die Gleichung (2.10) nicht mehr gilt. Deshalb wird Gleichung (2.10) zu (2.19 bzw. 2.20). Weit außerhalb der Verarmungszone gelten aber immer noch die gleichen Differenzen zwischen jetzt Quasi-Fermi Energien und intrinsischer Fermi Energie. Im Sperrbetrieb *erhöht* sich also die Spannung, die über der Verarmungsschicht abfällt gerade um die von außen angelegte Spannung. Ist die Diode in Durchlaßrichtung gepolt, so *verringert* sich die Potentialdifferenz der p- und n-Zone um den Betrag der äußeren Spannung.

## 2.3 Beweglichkeit

Im vorhergehenden Abschnitt wurden die Potentiale erklärt, die zur Lösung der Poisson-Gleichung (2.1) nötig sind. Aufgrund des berechneten intrinsischen Potentials werden die Ladungsträger je nach ihrem Vorzeichen beschleunigt, was einen Stromfluß zur Folge hat. Die Ladungsträger bewegen sich aber nicht frei im Festkörper, sondern unterliegen verschiedenen *Streumechanismen*, die ihre Bewegung abbremsen. Die wichtigsten Ursachen dafür sind:

- Gitterschwingungen (Phononen)
- Verunreinigungen
- die Oberfläche des Halbleiters
- andere Ladungsträger
- Gitterfehler aller Art

Um die mikroskopischen Ursachen durch eine makroskopische Größe beschreiben zu können führt man wie in den Gleichungen (2.21 und 2.22) die *Beweglichkeit*  $\mu_{n,p}$  der Elektronen bzw. Löcher ein. In der Beweglichkeit stecken dann die mikroskopischen Ursachen der Streuung der Ladungsträger. Sie wird dort als Konstante eingeführt, hängt aber wie oben gezeigt von verschiedenen Ursachen ab, so zum Beispiel von der Höhe der Dotierung, die nicht im kompletten Halbleiter gleich ist ( $\mu = \mu(x, y, z, \dots)$ ). Um die Beweglichkeit der Ladungsträger zu beschreiben, werden verschiedene Modelle für verschiedene Bauteile mit unterschiedlichen physikalischen Bedingungen benutzt. Will man zum Beispiel einen MOSFET<sup>4</sup> simulieren, so muß beachtet werden, daß sich die aktive Zone einer MOS-Struktur nur einige Nanometer unter der Oberfläche des Halbleiterkristalles befindet. In diesem Fall werden Streumechanismen der Ladungsträger an der Oberfläche eine Rolle spielen. Simuliert man dagegen eine Struktur, deren pn-Übergangsschicht einige Micrometer tief im Substrat liegt, können Effekte der Oberfläche vernachlässigt werden. Man sieht, daß die oben abgesteckten Grenzen zwischen Bauelementsimulationen in Schaltkreisen und den Device-Simulationen nicht so scharf

<sup>4</sup>MOSFET: Metal-Oxide-Semiconductor Field Effect Transistor



zu trennen sind. Auch hier müssen empirische Daten in die Modelle einfließen, und man muß je nach Situation entscheiden, welches Modell den gegebenen Umständen am besten entspricht. Desweiteren ist man auch hier daran interessiert, die Rechenzeit möglichst gering zu halten, so daß man sich auf die wesentlichen Einflüsse bezüglich der Beweglichkeit der Ladungsträger beschränkt. Es werden jetzt jeweils zwei Modelle, die in der vorliegenden Arbeit benutzt wurden, für ein Bipolar-Bauteil und eine MOSFET-Struktur gezeigt.

Beweglichkeitsmodelle in Bipolar-Strukturen:

Die einfachste Version besteht darin, für einen Halbleiter mit einem gewissen Dotierprofil die Beweglichkeiten für die Elektronen und Löcher in Form einer Tabelle anzugeben. Diese Tabelle wird von den Herstellern des Simulators mitgeliefert und beinhaltet Standardwerte der Mobilität in Abhängigkeit der Dotierkonzentrationen (siehe Tabelle 2.1). Sie berücksichtigt also die ersten beiden oben erwähnten Streumechanismen, nämlich die Streuung durch Phononen (die Tabelle gilt bei  $300K$ ) und die Streuung an Verunreinigungen (Dotierkonzentration).

Diese Modellierung der Beweglichkeiten ist gültig, solange die beschleunigenden Felder *klein* sind. Die Felder dürfen also eine gewisse Feldstärke nicht überschreiten. Betrachtet man einen Halbleiter bei Raumtemperatur und faßt die Elektronen als klassische Teilchen auf, so kann man eine Abschätzung der kritischen Feldstärke erhalten. Gemäß der Beziehung  $|\vec{v}| = \sqrt{\frac{2E}{m}}$  besitzen die Elektronen bei Raumtemperatur eine *thermische* Geschwindigkeit von  $10^6 \frac{cm}{s} \leq v_{th} \leq 10^7 \frac{cm}{s}$ , die ungeordnet in alle Richtungen verteilt ist. Legt man ein elektrisches Feld an, so überlagert sich der thermischen Geschwindigkeit  $v_{th}$  der Elektronen die Driftgeschwindigkeit  $v_d$ . Kleine Felder sind nun solche, bei denen  $v_d < v_{th}$  ist. In diesem Bereich gilt  $|\vec{v}| \propto |\vec{E}|$  mit der Proportionalitätskonstanten  $\mu$ . Die zugehörigen kritischen Feldstärken betragen  $|\vec{E}_{krit}| \approx 10^4 \frac{V}{cm}$ . Die Beweglichkeiten nehmen dann aufgrund zusätzlich eintretender Streuprozesse ab, so daß die ohm'sche Beziehung nicht mehr gilt. Im Simulator wird dies durch eine *feldabhängige Beweglichkeit* (dort mit FLDMOB abgekürzt) berücksichtigt. Das folgende Modell ist daher auch für die sogenannten großen Felder brauchbar, bei denen den Ladungsträgern so viel Energie zugeführt wird, daß deren Driftgeschwindigkeit in den Bereich der thermischen Geschwindigkeiten gelangt.

FLDMOB:

$$\mu_{n,p}(E) = \mu_{0\ n,p} \left[ \frac{1}{1 + \left( \frac{\mu_{0\ n,p} E}{v_{sat\ n,p}} \right)^{\beta_{n,p}}} \right]^{\frac{1}{\beta_{n,p}}} \quad (2.23)$$

mit der Sättigungsgeschwindigkeit  $v_{sat\ n,p}$ :

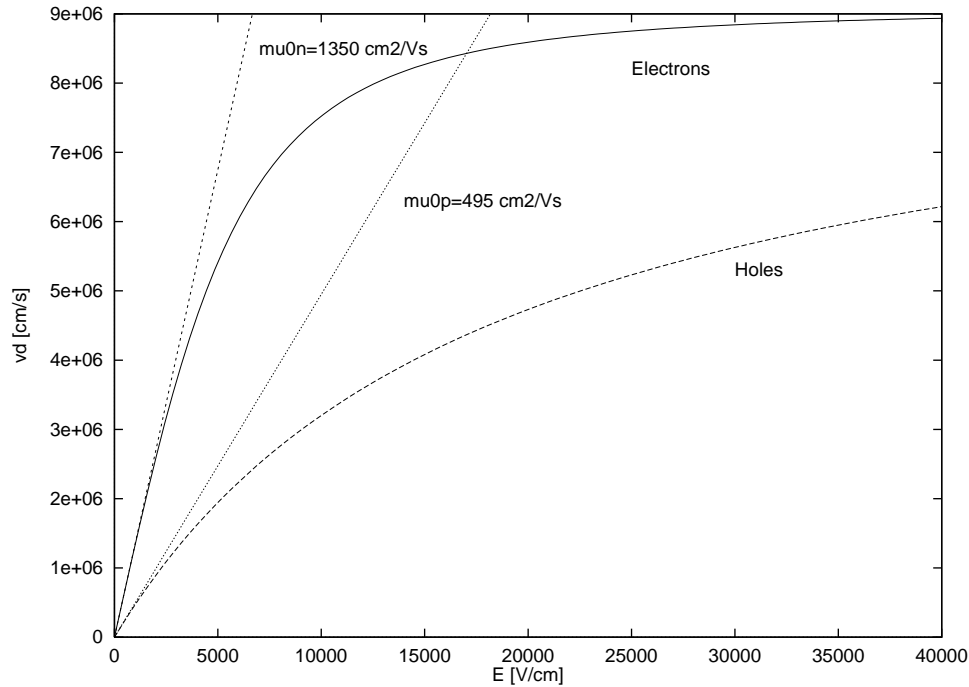
$$v_{sat\ n,p}(T_L) = \frac{2.4 \cdot 10^7}{1 + 0.8 \exp \left[ \frac{T_L}{600} \right]} \quad (2.24)$$

Dabei bedeuten:

$E$  den Betrag der Feldstärke,

Konzentration [ $cm^{-3}$ ]	Mobilität [ $cm^2/(Vs)$ ]	
	Elektronen	Löcher
$1 \cdot 10^{14}$	1350.0	495.0
$2 \cdot 10^{14}$	1345.0	495.0
$4 \cdot 10^{14}$	1335.0	495.0
$6 \cdot 10^{14}$	1320.0	495.0
$8 \cdot 10^{14}$	1310.0	495.0
$1 \cdot 10^{15}$	1300.0	491.1
$2 \cdot 10^{15}$	1248.0	487.3
$4 \cdot 10^{15}$	1200.0	480.1
$6 \cdot 10^{15}$	1156.0	473.3
$8 \cdot 10^{15}$	1115.0	466.9
$1 \cdot 10^{16}$	1076.0	460.9
$2 \cdot 10^{16}$	960.0	434.8
$4 \cdot 10^{16}$	845.0	396.5
$6 \cdot 10^{16}$	760.0	369.2
$8 \cdot 10^{16}$	720.0	348.3
$1 \cdot 10^{17}$	675.0	331.5
$2 \cdot 10^{17}$	524.0	279.0
$4 \cdot 10^{17}$	385.0	229.8
$6 \cdot 10^{17}$	321.0	203.8
$8 \cdot 10^{17}$	279.0	186.9
$1 \cdot 10^{18}$	252.0	178.0
$2 \cdot 10^{18}$	182.5	130.0
$4 \cdot 10^{18}$	140.6	90.0
$6 \cdot 10^{18}$	113.6	74.5
$8 \cdot 10^{18}$	99.5	66.6
$1 \cdot 10^{19}$	90.5	61.0
$2 \cdot 10^{19}$	86.9	55.0
$4 \cdot 10^{19}$	83.4	53.7
$6 \cdot 10^{19}$	78.8	52.9
$8 \cdot 10^{19}$	71.6	52.4
$1 \cdot 10^{20}$	67.8	52.0
$1 \cdot 10^{20}$	52.0	50.8
$1 \cdot 10^{20}$	35.5	49.6
$1 \cdot 10^{20}$	23.6	48.9
$1 \cdot 10^{20}$	19.0	48.4
$1 \cdot 10^{21}$	17.8	48.0

**Tabelle 2.1:** CONMOB: Beweglichkeit von Elektronen und Löchern in Silizium bei 300K (aus [8])



**Abbildung 2.4:** Driftgeschwindigkeit der Ladungsträger im FLDMOB-Modell (Erklärung siehe Text)

$\mu_{0\ n,p}$  die dotierkonzentrationsabhängige Beweglichkeit der Elektronen bzw. Löcher nach Tabelle 2.1,

$\beta_{n,p}$  ein Parameter, der standardmäßig für Elektronen 2 und für Löcher 1 ist.

Die Beweglichkeit sinkt also mit steigendem Feld. In der Abbildung 2.4 sieht man den Verlauf der Driftgeschwindigkeit  $v_d = \mu|\vec{E}|$  mit Gleichung (2.23) für die Beweglichkeiten von Elektronen und Löcher in Abhängigkeit des beschleunigenden Feldes. Es wurde hier von einer Anfangsbeweglichkeit der Elektronen von  $\mu_{0\ n} = 1350\text{cm}^2/\text{Vs}$  und der Löcher von  $\mu_{0\ p} = 495\text{cm}^2/\text{Vs}$  ausgegangen. Das entspricht nach Tabelle 2.1 einer Dotierkonzentration von  $10^{14}\text{Atome}/\text{cm}^3$ . Zusätzlich sind die beiden Geraden  $\mu_{0\ n}|\vec{E}|$  und  $\mu_{0\ p}|\vec{E}|$  eingezeichnet. Dies wäre der Verlauf der Driftgeschwindigkeiten, falls die Mobilitäten der Ladungsträger unabhängig von den beschleunigenden Feldern wären. Man sieht, daß dies nur für sehr kleine Felder gelten kann.

Beweglichkeiten in MOSFETs:

In MOSFET-Tansistoren bilden sich, bei genügend hoher Gate-Spannung, unterhalb des Gates nahe der Oberfläche des Halbleiterchips sogenannte *inversion layers* aus. Diese sind die aktiven Zonen eines MOSFETs. Um die Bewegung von Ladungsträgern in Inversionsschichten zu beschreiben, bzw. die Streuprozesse nahe der Oberfläche des Halbleiterkristalls zu berücksichtigen, dient als Parameter das *transversale elektrische Feld*  $E_{\perp}$ .

Das sogenannte CVT-Modell besteht aus drei Anteilen, aus denen sich die endgültige Beweglichkeit zusammensetzt. Es werden drei Mechanismen unterschieden, die die Mobilität der Ladungsträger an der Oberfläche begrenzen:

1.  $\mu_{AC}$ : Streuung an akustischen Phononen

$$\mu_{AC} = \frac{\left[ \frac{BT}{E_{\perp}} + \frac{CN^{\tau}}{E_{\perp}^{1/3}} \right]}{T}$$

2.  $\mu_{sr}$ : Streuung durch Unebenheiten der Oberfläche

$$\mu_{sr} = \frac{\delta}{E_{\perp}^2}$$

3.  $\mu_b$ : Streuung an optischen Phononen

$$\mu_b = \mu_0 e^{-\frac{P_c}{N}} + \frac{\left[ \mu_{max} \left( \frac{T}{300} \right)^{-\gamma} - \mu_0 \right]}{1 + \left( \frac{N}{C_r} \right)^{\alpha}} - \frac{\mu_1}{1 + \left( \frac{C_s}{N} \right)^{\beta}}$$

$N$  ist die Gesamtdichte der Dotierung. In der Tabelle 2.2 sind die benutzten Parameter aufgelistet.

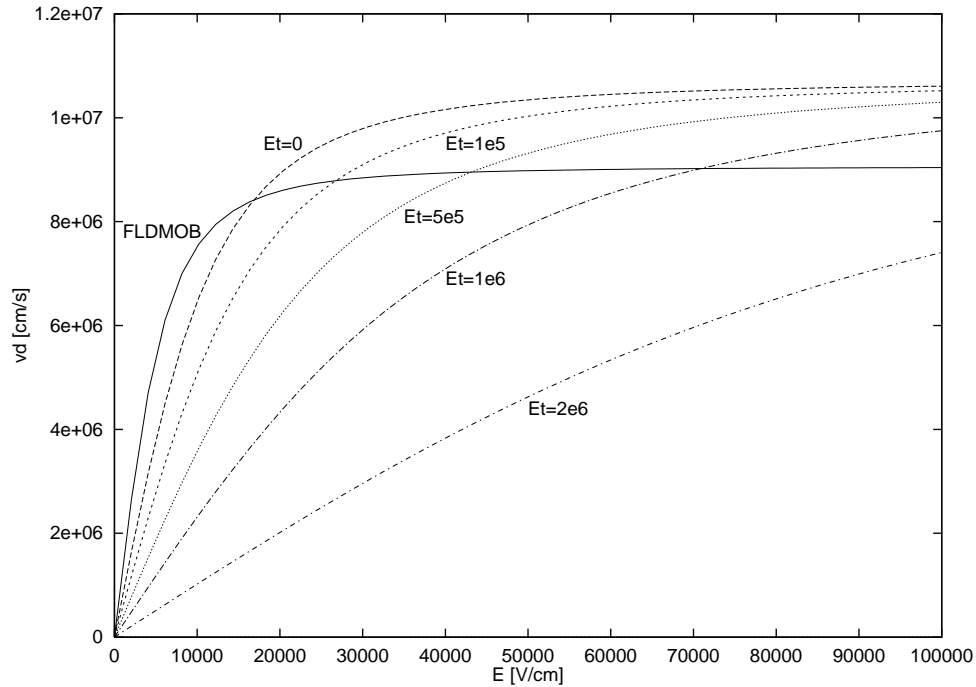
Parameter	Wert		Einheit
	Elektronen	Löcher	
$B$	$4.75 \cdot 10^7$	$9.925 \cdot 10^6$	$cm/s$
$C$	$1.74 \cdot 10^5$	$8.842 \cdot 10^5$	$cm^{5/3} K / (V^{2/3} s)$
$\tau$	0.125	0.0317	-
$\gamma$	2.5	2.2	-
$\mu_0$	52.2	44.9	$cm^2 / (Vs)$
$\mu_1$	43.4	29.0	$cm^2 / (Vs)$
$\mu_{max}$	1417.0	470.5	$cm^2 / (Vs)$
$C_r$	$9.68 \cdot 10^{16}$	$2.23 \cdot 10^{17}$	$cm^{-3}$
$C_s$	$3.43 \cdot 10^{20}$	$6.10 \cdot 10^{20}$	$cm^{-3}$
$\alpha$	0.680	0.71	-
$\beta$	2.0	2.0	-
$P_c$	0.0	$0.23 \cdot 10^{16}$	$cm^{-3}$
$\delta$	$5.82 \cdot 10^{14}$	$2.0546 \cdot 10^{14}$	$V/s$

**Tabelle 2.2:** Standardparameter im CVT-Modell (aus [8]).

Die drei Anteile werden gemäß der *Mathies'schen Regel* zur Gesamtmobilität zusammengesetzt. Die Mathies'sche Regel besagt unter der Voraussetzung der Unabhängigkeit der einzelnen Streumechanismen, daß der kleinste Anteil der Beweglichkeit die Gesamtbeweglichkeit begrenzt:

$$\frac{1}{\mu_T} = \frac{1}{\mu_{AC}} + \frac{1}{\mu_{sr}} + \frac{1}{\mu_b}$$

Die endgültige Formel für die *parallelfeldabhängige* Beweglichkeit im CVT-Modell sieht derjenigen im FLDMOB-Modell von oben sehr ähnlich, mit den Unterschieden, daß  $\mu_0$  aus Gleichung (2.23) durch die resultierende transversalfeldabhängige Beweglichkeit  $\mu_T$  ersetzt wird, und daß die Sättigungsgeschwindigkeit  $v_{sat}$  etwas anders berechnet wird.



**Abbildung 2.5:** Driftgeschwindigkeit der Elektronen im CVT-Modell (Erklärung siehe Text)

CVT:

$$\mu_{n,p} = \mu_{T n,p} \left[ \frac{1}{1 + \left( \frac{\mu_{T n,p} E}{v_{sat n,p} \left( \frac{T}{300} \right)^\lambda} \right)^{\beta_{n,p}}} \right]^{\frac{1}{\beta_{n,p}}} \quad (2.25)$$

Die Parameter folgen in der Tabelle 2.3.

Parameter	Wert		Einheit
	Elektronen	Löcher	
$v_{sat}$	$10.7 \cdot 10^6$	$8.34 \cdot 10^6$	$cm/s$
$\lambda$	-0.87	-0.52	-
$\beta$	2.0	1.0	-

**Tabelle 2.3:** Parameter für die Gleichung (2.25) (aus [8]).

In der Abbildung 2.5 ist die Driftgeschwindigkeit  $v_d = \mu E_{\parallel}$  mit der Beweglichkeit des CVT-Modelles gezeigt. Die angenommene Dotierkonzentration beträgt  $7 \cdot 10^{16} cm^{-3}$ . Es sind hier zur Übersicht nur die Kurven für die Elektronen gezeigt. Als Parameter für die einzelnen Kurven dient das Transversalfeld  $E_{\perp}$  (im Bild als  $E_t$  bezeichnet). Zum Vergleich dazu ist die Driftgeschwindigkeit im FLDMOB-Modell mit eingezeichnet. Man sieht, wie die Driftgeschwindigkeit der Elektronen, aufgrund der abnehmenden Mobilität, mit zunehmendem

Transversalfeld  $E_{\perp}$  abnimmt.

Als letztes Modell für die Beweglichkeit der Ladungsträger wird das YAMAGUCHI-Modell vorgestellt. Es berücksichtigt ebenfalls eine Transversalfeld-Abhängigkeit. Das Modell besteht aus zwei Teilen:

1. Kleine longitudinale Felder ( $E_{\parallel}$ ) und kein transversales Feld:

$$\mu_0 = \mu_L \left[ 1 + \frac{N_i}{\left(\frac{N_i}{S} + N_{ref}\right)} \right]^{-1/2}$$

2. Transversalfeld-Abhängigkeit:

$$\mu_s = \mu_0(1 + \alpha_s E_{\perp})^{-1/2}$$

Die endgültige Gleichung für die Beweglichkeit im YAMAGUCHI-Modell lautet:

YAMAGUCHI:

$$\mu = \mu_s \left[ 1 + \left( \frac{\mu_s E_{\parallel}}{u_l} \right)^2 \left( \frac{G + \mu_s E_{\parallel}}{u_l} \right)^{-1} + \left( \frac{\mu_s E_{\parallel}}{v_{sat}} \right)^2 \right]^{-1/2} \quad (2.26)$$

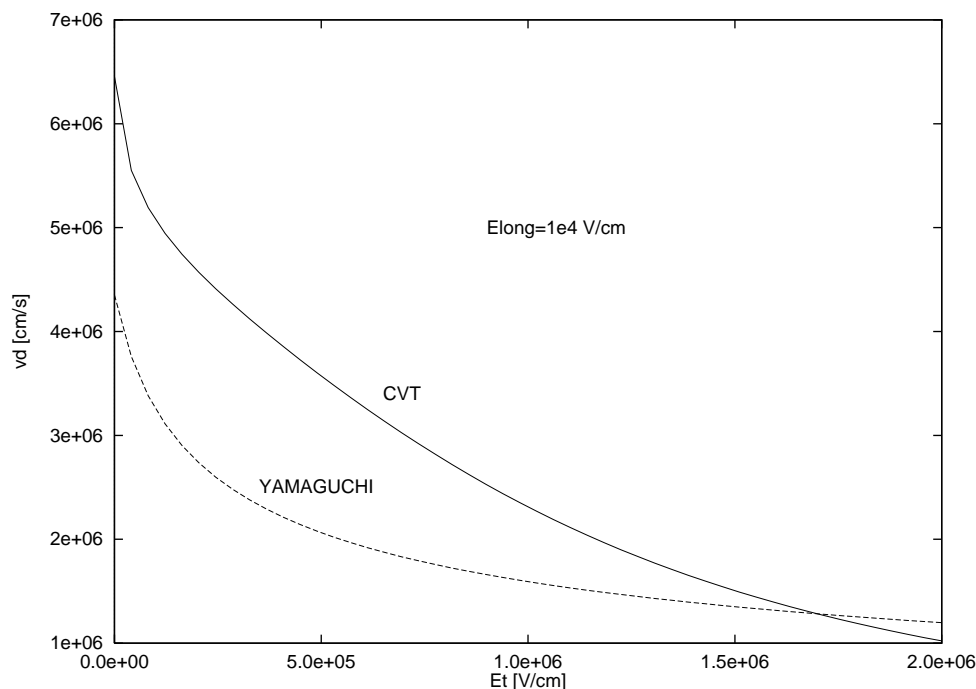
Die Parameter folgen in der Tabelle 2.4.

Parameter	Wert		Einheit
	Elektronen	Löcher	
$S$	350.0	81.0	-
$N_{ref}$	$3.0 \cdot 10^{16}$	$4.0 \cdot 10^{16}$	$cm^{-3}$
$\mu_L$	1400.0	480.0	$cm^2/(Vs)$
$\alpha_s$	$1.54 \cdot 10^{-5}$	$5.35 \cdot 10^{-5}$	$cm/V$
$v_{sat}$	$1.036 \cdot 10^7$	$1.2 \cdot 10^7$	$cm/s$
$u_l$	$4.9 \cdot 10^6$	$2.92 \cdot 10^6$	$cm/s$
$G$	8.8	1.6	$cm/s$

**Tabelle 2.4:** Parameter für die Gleichung (2.26) (aus [8]).

Um zu verdeutlichen, wie das Transversalfeld die Driftgeschwindigkeiten im CVT- und YAMAGUCHI-Modell bestimmt, ist in der Abbildung 2.6 die Driftgeschwindigkeit bei einer Dotierkonzentration von  $7 \cdot 10^{16} cm^{-3}$  in Abhängigkeit des *transversalen* Feldes bei festgehaltenem longitudinalem Feld geplottet.

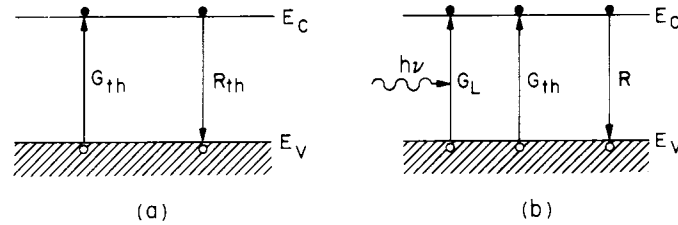
Die beiden Modelle zeigen zwar etwas unterschiedliches Verhalten bezüglich der  $E_{\perp}$ -Abhängigkeit, die Unterschiede in den Ausgangs- und Übertragungskennlinien eines simulierten MOS-FETs sind aber gering.



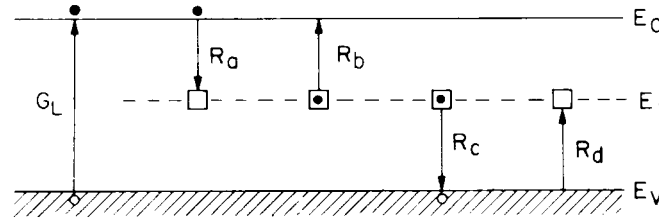
**Abbildung 2.6:** Vergleich der Modelle CVT-YAMAGUCHI in Abhängigkeit des transversalen Feldes bei festem longitudinalem Feld

## 2.4 Generation und Rekombination von Ladungsträgern

Befindet sich ein Halbleiter nicht im thermischen Gleichgewicht, so ist er bestrebt, dahin zurück zu gelangen. Die Größe, die die Abweichung vom thermischen Gleichgewicht beschreibt, ist das Produkt der Ladungsträgerkonzentrationen  $n \cdot p$ . Im Gleichgewicht gilt das Massenwirkungsgesetz  $n \cdot p = n_i^2$ . Ist das Produkt  $n \cdot p < n_i^2$ , so spricht man von *Extraktion*, da der Halbleiter nun weniger freie Ladungsträger besitzt als im Gleichgewichtszustand (das ist zum Beispiel bei einer in Sperrichtung gepolten Diode der Fall). Im anderen Falle gilt  $n \cdot p > n_i^2$ . Dieser wird als *Injektion* bezeichnet. Im Halbleiter gibt es nun zwei Mechanismen, die für das Wiedererlangen des Gleichgewichtszustandes verantwortlich sind. Das ist im Falle der Extraktion die Generation von Ladungsträgern bzw. bei der Injektion die Rekombination derer. Weil das thermische Gleichgewicht kein statischer Zustand, sondern ein dynamischer Prozeß ist, werden in einem Gleichgewichtszustand permanent Ladungsträgern generiert und rekombiniert. Die Generationsrate  $G$  und Rekombinationsrate  $R$  müssen dann aber gleich groß sein. (Die Raten  $G$  und  $R$  sind in einem *stationären* Nichtgleichgewichtszustand (z.B. konstante Lichteinstrahlung) ebenfalls gleich ( $R = G$ ), aber im Gegensatz zum thermischen Gleichgewicht gilt:  $n \cdot p \neq n_i^2$ .) Betrachtet man den Extraktionsfall, so muß, um zum Gleichgewicht zurückzukehren, die Generationsrate größer sein als die Rekombinationsrate. Entsprechend muß nach einer Ladungsträgerinjektion die Rekombinationsrate überwiegen. Weiter oben wurden bereits Situationen beschrieben, die eine Abweichung vom Gleichgewicht zeigten. In den Gleichungen (2.19 und 2.20) wurden von außen angelegte Spannungen (Quasi-Fermi Levels) dadurch berücksichtigt, daß sie je nach Vorzeichen und Größe das Produkt  $n \cdot p$  gegenüber  $n_i^2$  vergrößern oder verkleinern. Eine weitere Möglichkeit



**Abbildung 2.7:** Direkte Generation und Rekombination im thermischen Gleichgewicht (a) und unter Lichteinstrahlung (b) (aus [6])



**Abbildung 2.8:** Indirekte Generation und Rekombination unter Lichteinstrahlung (aus [6])

Überschußladungen zu produzieren, ist Licht geeigneter Wellenlänge auf einen Halbleiter zu scheinen. Ist die Energie eines einfallenden Photons größer als die Bandlücke im Halbleiter, so kann ein Elektron über die Bandlücke vom Valenz- ins Leitungsband gehoben werden.

Es gibt verschiedene Möglichkeiten für einen Halbleiter, zum Gleichgewichtszustand zurückzukehren. Man unterscheidet *direkte* und *indirekte* Rekombination. Bei der direkten Rekombination fällt ein vormals angeregtes Elektron im Leitungsband direkt ins Valenzband und gibt seine Energie ab. Bei indirekten Halbleitern befinden sich das Maximum der Valenzbandkante und das Minimum der Leitungsbandkante nicht an der gleichen Stelle im  $\vec{k}$ -Raum, so daß die beiden Zustände unterschiedliche Impulszustände besitzen. Das emittierte Photon kann dem Elektron zwar Energie entziehen, ist aber aufgrund seines geringen Impulses ( $p_{Ph} = \hbar/\lambda$ ) nicht in der Lage, die nötige  $\vec{k}$ -Differenz zu liefern. Die Impulsbilanz muß durch ein mitwirkendes Phonon ausgeglichen werden, so daß bei indirekten Halbleitern die Band-Band Rekombination sehr unwahrscheinlich wird.

Bei indirekten Halbleitern überwiegen daher Rekombinationen über sogenannte *single level*-Prozesse mit Hilfe von Rekombinationszentren, die ungefähr in der Mitte der verbotenen Zone liegen. Die Energie des Elektrons wird dann nicht als Strahlung emittiert, sondern überträgt sich auf das Gitter, das heißt, es werden entsprechende Phononen angeregt. Bei direkten Halbleitern (z.B. GaAs) befinden sich das Maximum des Valenzbandes und das Minimum des Leitungsbandes an der selben Stelle des  $\vec{k}$ -Vektors, so daß es keine Impulsdifferenz der beiden Zustände gibt und somit einer Band-Band Rekombination nichts im Wege steht. Aber auch bei direkten Halbleitern kann ein angeregtes Elektron im Leitungsband strahlungslos seine Energie abgeben, indem es seine Energie an ein weiteres Elektron oder Loch überträgt. Dieser Prozeß wird als *Auger*-Rekombination bezeichnet. In den Abbildungen 2.7 und 2.8 sind jeweils direkte und indirekte Rekombinationsprozesse zu sehen. Die Statistik der Rekombina-



tionszentren wird nach W. Shockley, W.T. Read und R.N. Hall (SRH-Statistik) beschrieben.

SRH-Rekombination:

$$R_{SRH} = \frac{n \cdot p - n_i^2}{\tau_p \left[ n + n_{ie} e^{\left(\frac{\epsilon_t - \epsilon_i}{kT_L}\right)} \right] + \tau_n \left[ p + n_{ie} e^{\left(\frac{\epsilon_i - \epsilon_t}{kT_L}\right)} \right]} \quad (2.27)$$

Die Lebensdauern der Ladungsträger werden dotierkonzentrationsabhängig modelliert:

$$\tau_n = \frac{\tau_{n0}}{1 + N/N_{SRH_n}}$$

bzw.

$$\tau_p = \frac{\tau_{p0}}{1 + N/N_{SRH_p}},$$

dabei ist N die lokale absolute Dotierkonzentration  $N_D + N_A$ . Die weiteren Parameter folgen in der Tabelle 2.5.

Parameter	Wert	Einheit
$\epsilon_t - \epsilon_i$	0.0	eV
$\tau_{n0}$	$1 \cdot 10^{-7}$	s
$N_{SRH_n}$	$5 \cdot 10^{16}$	$1/cm^3$
$\tau_{p0}$	$1 \cdot 10^{-7}$	s
$N_{SRH_p}$	$5 \cdot 10^{16}$	$1/cm^3$

**Tabelle 2.5:** Parameter für die Gleichung (2.27) (aus [8]).

Zur Vollständigkeit wird die strahlende Rekombination aufgeführt, die aber hier keine Rolle spielt, da ausschließlich Silizium Halbleiter betrachtet werden.

Strahlende Rekombination:

$$R_{optical} = C_{opt}(p \cdot n - n_{ie}^2)$$

Der Parameter  $C_{opt}$  ist wie oben erwähnt zu null gesetzt.

Auger-Rekombination:

$$R_{Auger} = c_n (p \cdot n^2 - n \cdot n_{ie}^2) + c_p (n \cdot p^2 - p \cdot n_{ie}^2) \quad (2.28)$$

mit den Parametern:

An der Oberfläche eines Halbleiterkristalls befinden sich aufgrund der nicht abgesättigten Bindungen der obersten Atomlage zusätzliche Rekombinationszentren, die die Rekombinationsraten erhöhen. Um Effekte bei Bauteilen, deren aktive Zone sich nahe der Oberfläche befindet (z.B. MOSFET's), zu berücksichtigen, wird die Oberflächenrekombination durch eine Flächenstromdichte modelliert.

Parameter	Wert	Einheit
$c_n$	$8.3 \cdot 10^{-32}$	$cm^6/s$
$c_p$	$1.8 \cdot 10^{-31}$	$cm^6/s$

**Tabelle 2.6:** Parameter für die Gleichung (2.28) (aus [8])

Oberflächenrekombination:

$$J_{sn} = qv_{sn}(n_s - n_{eq})$$

bzw. für die Löcher

$$J_{sp} = qv_{sp}(p_s - p_{eq})$$

Die Parameter  $v_{sn}$  und  $v_{sp}$  können vom Benutzer eingestellt werden.

Alle in diesem Kapitel vorgestellten Formeln sind im *Atlas User's Manual, Device Simulation Software, Chapter 3* der Firma *Silvaco International* aufgeführt ([8]).

## Kapitel 3

# Der pn-Übergang

Als pn-Übergang bezeichnet man die Grenze zwischen einem n-dotierten und einem p-dotierten Gebiet in einem Halbleiter. Gedanklich kann man sich vorstellen, daß ein n-dotiertes Halbleiterplättchen mit einem p-dotiertem in Kontakt gebracht wird. Aufgrund der unterschiedlichen Dotierungsart in p- und n-Gebiet existieren im p-Gebiet sehr viel mehr freie p-Ladungsträger (im folgenden als *Löcher*) bezeichnet als n-Ladungsträger (*Elektronen*). Analog dazu befinden sich im n-dotierten Halbleiter sehr viel mehr Elektronen als Löcher. Stoßen nun die beiden Gebiete aufeinander, und damit eine sehr hohe Löcherkonzentration im p-Halbleiter auf eine sehr geringe Löcherkonzentration im n-Halbleiter, bzw. umgekehrt eine sehr hohe Elektronenkonzentration im n-Gebiet auf eine sehr geringe im p-Gebiet, so versuchen sich die Konzentrationen durch Diffusion auszugleichen. Diese Wanderung der Ladungsträger – also die der Elektronen vom n- ins p-Gebiet und die der Löcher vom p- ins n-Gebiet – bewirkt einen elektrischen Strom in der Nähe der Übergangszone und wird der Entstehung entsprechend als Diffusionsstrom bezeichnet. Wenn nun die Elektronen vom n- ins p-Gebiet diffundiert sind, so sind sie dort Minoritäten und rekombinieren dort schnell mit den Löchern, die dort Majoritäten sind. Das heißt, die ins p-Gebiet diffundierten Elektronen besetzen die Akzeptoren der p-Dotierung und vernichten damit die „freien“ Löcher. Analoge Betrachtungen gelten für die Löcher, die auf die n-Seite diffundieren. Dadurch wird aber die lokale Ladungsneutralität aufgehoben, denn die Akzeptoren der p-Seite werden durch den Einfang von Elektronen negativ geladen, ebenso wie die Donatoren des n-Gebietes positiv geladen werden. Es entsteht eine *Raumladungszone* (aufgrund des Fehlens freier Ladungsträger in diesem Bereich auch *Verarmungszone* (engl.: depletion) genannt) im wesentlichen bestehend aus den ionisierten, im Gitter eingebauten Störatomen<sup>1</sup>. In Kap.1 wurde erläutert, daß die Donatoren bei Zimmertemperatur ionisiert sind, also ihr überschüssiges Elektron schon ins Leitungsband abgegeben haben. Dadurch wird aber die räumliche Ladungsneutralität nicht gestört, denn „ins Leitungsband wechseln“ bedeutet ja nur, ein anderes Energieniveau einzunehmen, wogegen sich durch den Prozeß der Diffusion eine *örtliche* Fluktuation der Ladungsträger ergibt. Die ionisierten Störatome beiderseits des pn-Überganges bewirken nun ein elektrisches Feld, welches vom n- zum p-Gebiet gerichtet ist. Dieses Feld bewirkt nun wiederum eine Bewegung der Ladungsträger längs der Feldlinien, wobei die Löcher in Richtung der Feldlinien beschleunigt werden und die Elektronen diesen entgegen. Dieser Strom, bestehend aus Elektronen

---

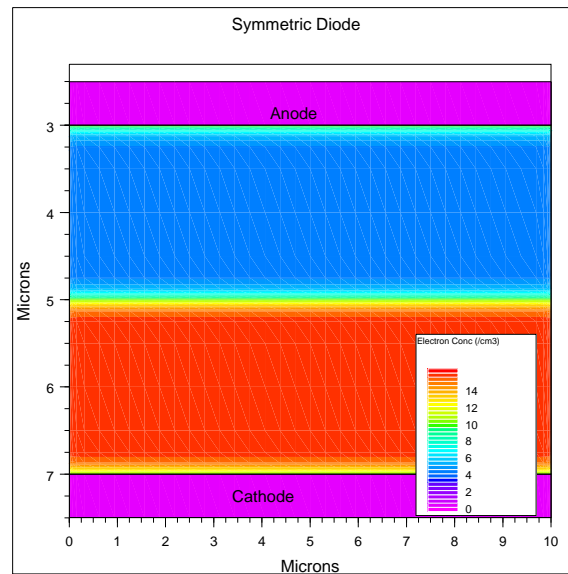
<sup>1</sup>Die über den Übergang diffundierten Ladungsträger tragen auch zur Raumladungszone bei, sind aber, aufgrund der Rekombination, in ihren Dichten weit geringer als die ionisierten Störatome.

und Löchern in entgegengesetzten Ausbreitungsrichtungen, wird *Feldstrom* genannt. Der Feldstrom wirkt dem Diffusionsstrom entgegen, so daß sich ein dynamisches Gleichgewicht ausbildet. Befindet sich der Halbleiter pn-Übergang im thermodynamischen Gleichgewicht, so muß gelten:

$$|I_{diff}| = |I_{feld}|$$

### 3.1 Symmetrisch abrupt dotierte Diode

Das einfachste Halbleiterbauelement, die Diode, besteht aus genau einem pn-Übergang. In Abbildung 3.1 ist die Elektronenkonzentration einer symmetrisch und abrupt dotierten Diode im thermodynamischen Gleichgewicht gezeigt.



**Abbildung 3.1:** Symmetrisch abrupt dotierte Diode mit jeweils  $10^{16}/cm^3$  n- (Phosphor) und p- (Bor) dotiert

Die Farbskala zeigt die Elektronenkonzentration der Diode an. In der unteren Hälfte des Bildes ist das n-Gebiet – mit dem Aluminiumkontakt als Kathode – zu sehen, das mit  $10^{16}/cm^3$  dotiert ist. Im folgenden dienen als Donatoren, falls nicht speziell angegeben, immer Phosphoratome und als Akzeptoren Boratome. Wie vorher schon erläutert, sind bei Zimmertemperatur nahezu alle Phosphoratome ionisiert, so daß die Elektronenkonzentration im n-Gebiet ebenfalls  $10^{16}/cm^3$  beträgt. Im p-Gebiet, das mit ebenso vielen Boratomen dotiert ist, existieren, wie im Bild zu erkennen ist, praktisch keine Elektronen, denn aufgrund des Massenwirkungsgesetzes  $n \cdot p = n_i^2$ , folgt  $n = n_i^2/p \approx 10^{20}/(10^{16} cm^3) = 10^4/cm^3$ , was im Vergleich zu den Löchern einer verschwindend kleinen Konzentration entspricht. Das gilt analog für die Löcher auf der n-Seite.

Um einen Einblick in die Raumladungsdichten zu erhalten, wurde in der Abbildung 3.1 ein Schnitt (engl.: cutline) längs der Diode senkrecht zum pn-Übergang definiert. Die Abbildung

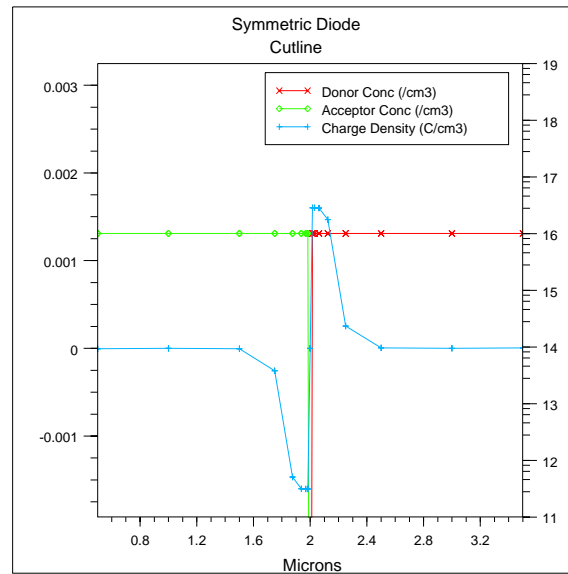


Abbildung 3.2: Schnitt durch die Verarmungszone

3.2 zeigt diesen Schnitt. Zu beachten ist, daß die Längenskala um  $3\mu m$  verschoben ist und daher die Raumladungszone, die in Abbildung 3.1 an der Stelle  $y = 5\mu m$  zu finden war, nun in der Abbildung 3.2 bei  $2\mu m$  zu liegen kommt. Zusätzlich zur blau eingezeichneten Raumladungsdichte  $\rho = q(N_D + p - N_A - n)$  sind die Konzentrationen der Akzeptoren im p-Gebiet (grün) und die der Donatoren im n-Gebiet (rot) zu sehen. Der Übergang ist wie oben schon erwähnt symmetrisch (d.h. die Akzeptorkonzentration ist im p-Gebiet gleich groß der Donatorkonzentration im n-Gebiet) und abrupt. Durch die Geometrie und das gegebene Dotierprofil ist die Ausbreitung der Raumladungszone ebenfalls symmetrisch, sie erstreckt sich also in beide Gebiete gleichförmig. Die Dicke der Raumladungszone ist abhängig von den Dotierkonzentrationen und der angelegten Spannung. Im hier vorliegenden Fall der abrupt dotierten Diode kann die Breite der Raumladungszone  $s_0$  berechnet werden und ergibt sich bei spannungslosem Zustand zu:

$$s_0 = \sqrt{2 \frac{\epsilon U_D}{q} \frac{N_D + N_A}{N_D N_A}}. \quad (3.1)$$

Mit den speziellen Werten der hier vorliegenden Diode ( $N_D = N_A = 10^{16}/cm^3$ ,  $U_D \approx 0.7V$  und  $\epsilon = \epsilon_{Si}\epsilon_0 = 11.7 \cdot \epsilon_0$ ) ergibt sich die Breite zu ca.  $425nm$ . Legt man an den pn-Übergang eine externe Spannung an, so verändern sich die Energien der Ladungsträger und die Raumladungszone wird modifiziert. Wird zum Beispiel dem Kontakt der n-Zone (=Kathode) eine positive Spannung zugeführt, während die Anode auf Masse liegt, so kann man sich anschaulich vorstellen, daß die vom n- ins p-Gebiet diffundierten Elektronen aufgrund des nun existierenden Feldes noch weiter von der Übergangsschicht abgezogen werden, so daß man eine Vergrößerung der Raumladungszone erwartet. Tatsächlich hängt die Breite  $s$  der Raumladungszone von der außen angelegten Spannung ab. Diese muß dann zur Diffusionsspannung (vorzeichengerecht) addiert werden, so daß man für die Breite  $s$  unter Einwirkung

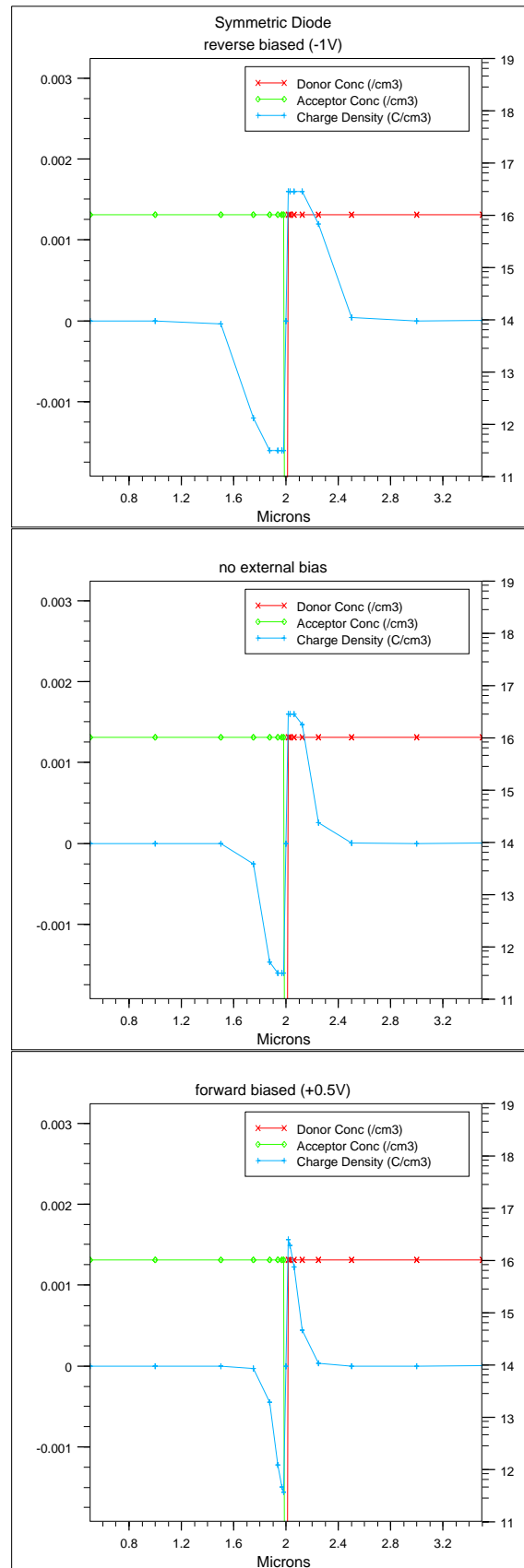
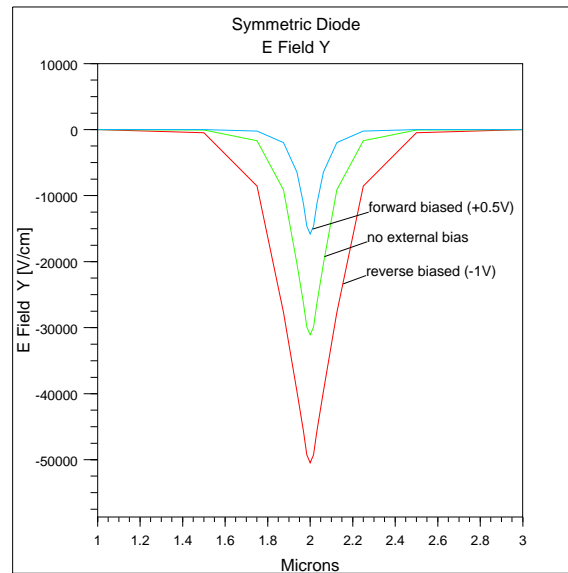


Abbildung 3.3: Vergleich der Raumladungszonen bei verschiedenen externen Spannungen

der Fremdspannung  $U_{fr}$  folgende Formel erhält:

$$s = \sqrt{2 \frac{\epsilon(U_D - U_{fr})}{q} \frac{N_D + N_A}{N_D N_A}} = s_0 \sqrt{1 - \frac{U_{fr}}{U_D}}, \quad (U_{fr} \leq U_D). \quad (3.2)$$

Hierbei wird die Fremdspannung positiv gezählt, wenn sie in Durchlaßrichtung (d.h. positiver Pol an der Anode und negativer Pol der Spannungsquelle an der Kathode) geschaltet wird und negativ in Sperrichtung. Sie ist also der Diffusionsspannung  $U_D$  entgegengerichtet. Für  $U_{fr} \rightarrow U_D$  wird die Raumladungszone immer schmaler und verschwindet schließlich, wenn  $U_{fr} = U_D$  ist. Die Modulation der Dicke der Raumladungszone durch die von außen angelegte Spannung wird auch als *Early-Effekt* bezeichnet. Die Abbildung 3.3 zeigt wieder den gleichen Schnitt und die gleichen Kurven, wobei hier aber von oben nach unten Spannungen von  $-1V$ ,  $0V$  und  $+0.5V$  an den pn-Übergang angelegt wurden. Man erkennt deutlich das Schrumpfen der Breite der Raumladungszone. Die Abbildung 3.4 verdeutlicht, ebenfalls längs der gleichen



**Abbildung 3.4:** Vergleich der y-Komponenten des elektrischen Feldes bei verschiedenen externen Spannungen

Schnittlinie, in einer Überlagerung dreier Kurven, die die y-Komponente des elektrischen Feldes<sup>2</sup> darstellen, wie gemäß der Modulation der Raumladungszone in Abhängigkeit der angelegten Spannung ebenfalls das elektrische Feld mit zunehmender Sperrspannung steigt.

Die Kennlinie, also der Anodenstrom in Abhängigkeit der Anodenspannung, kann ebenfalls mit den oben gemachten Idealisierungen näherungsweise berechnet werden. Mit den Annahmen der *Schottky-Näherung*, die besagen daß die Raumladungszone, nur von den festen ionisierten Störstellenatomen gebildet wird, die geringe Konzentration der diffundierten Minoritäten also vernachlässigt wird und die Störstellen innerhalb der Raumladungszone nahezu vollständig ionisiert sind (d.h.  $N_D^+ \approx N_D$  bzw.  $N_A^- \approx N_A$ ), ergibt sich ein rechteckiger

<sup>2</sup>Das ist diejenige Komponente, die senkrecht zur Verarmungszone verläuft

Ladungsdichtenverlauf in der pn-Übergangszone, die sich an der Stelle  $x = 0$  befinden soll:

$$\rho(\vec{x}) = \begin{cases} -qN_A & , \quad -s_0/2 < x < 0 \\ +qN_D & , \quad 0 < x < +s_0/2 \\ 0 & , \quad \text{sonst} \end{cases}$$

Mit dieser idealisierten Ladungsdichte ergibt sich die Stromdichte eines pn-Überganges unter Einwirkung einer äußeren Spannung  $U_{ex}$  zu:

$$J(U_{ex}) = J_s \left( e^{\frac{qU_{ex}}{kt}} - 1 \right) \quad (3.3)$$

$J_s$  ist die Sättigungsstromdichte, die über den pn-Übergang fließt, wenn dieser in Sperrichtung gepolt ist,  $U_{ex}$  also negativ wird. In Vorwärtsrichtung wächst der Strom mit der Spannung exponentiell an.

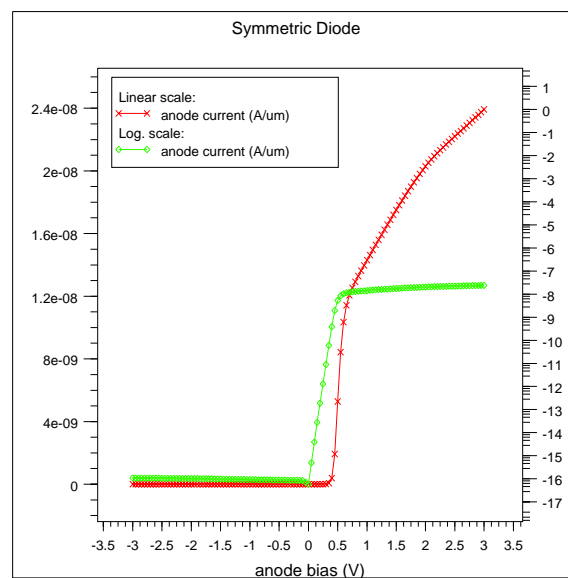


Abbildung 3.5: Kennlinie der Diode

In der Abbildung 3.5 ist die Kennlinie der obigen Diode gezeigt. Der Anodenstrom ist in Abhängigkeit der Anodenspannung linear (rot) und logarithmisch (grün) aufgetragen. Gemäß der Gleichung (3.3) fließt in Sperrichtung ( $U_{ex} < 0$ ) der nahezu spannungsunabhängige Sättigungsstrom. Wächst die Sperrspannung an, so verbreitert sich die Raumladungsschicht nach Gleichung (3.2) aber nur proportional der Quadratwurzel der Spannung, das heißt die Feldstärke über der Raumladungszone wächst an. Die Durchbruchspannung von Silizium beträgt ca.  $10^6 V/cm$ . Beim Erreichen dieser Feldstärke wächst der Rückwärtsstrom schlagartig an. Die Dotierung in diesem Falle ist aber zu gering, bzw. die Raumladungszone zu breit, als daß bei den hier vorliegenden Spannungen ein Durchbruch erfolgen könnte. An der Vorwärtskennlinie ist aber eine deutliche Abweichung von der Vorhersage der Gleichung (3.3) zu erkennen. An der Stelle  $U \approx 0.6V$  setzt sich der exponentielle Verlauf nicht mehr fort, wie in Gleichung (3.3), sondern die Kennlinie knickt ab, was besonders in der logarithmischen Darstellung zu sehen ist. Dies geschieht aufgrund von zusätzlichen Streuprozessen



der Ladungsträger untereinander, die bei zunehmender Konzentration immer mehr an Bedeutung gewinnen. Die Mobilität der Ladungsträger nimmt ab und die Driftgeschwindigkeit der Elektronen wächst nicht mehr linear mit dem wachsenden Feld an, sondern geht gegen eine Sättigungsgeschwindigkeit, wodurch der exponentielle Stromanstieg gebremst wird. Der Effekt wird auch als *Knie-Effekt* bezeichnet.

Für diesen idealisierten Fall ist es möglich, Formeln herzuleiten, die die Breite der Raumladungszone, die Felder darin usw. beschreiben. In der Praxis ist es aber unmöglich, solche Dioden herzustellen. Eine wesentliche Annahme, die in Gleichung (3.2) eingeht, ist die einer abrupten Dotierung. Das kann in der Realität nicht erreicht werden. In Wirklichkeit werden die verschiedenen Dotierkonzentrationen der verschiedenen Gebiete stetig ineinander übergehen.

Im nächsten Abschnitt werden Dioden besprochen, deren Dotierprofile nahezu der Realität entsprechen, bzw. deren Störstellenverteilung nach Herstellerangaben erzeugt wurde. Im ASIC-Labor Heidelberg werden von den Chip-Designern im wesentlichen drei verschiedene Dioden, die im folgenden nach ihrer Schichtfolge benannt werden, benutzt:

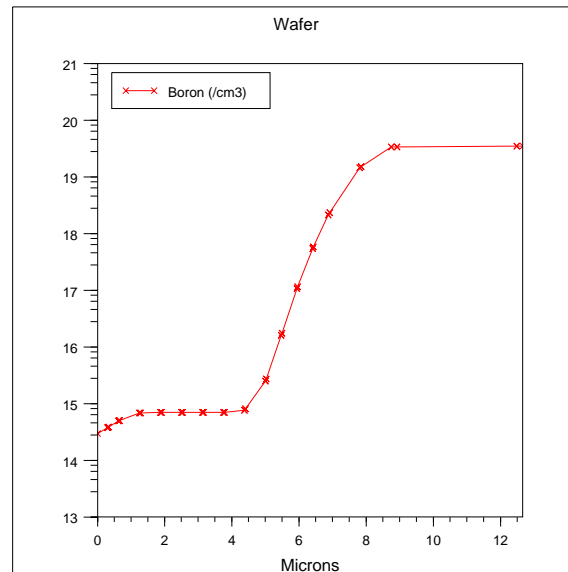
- N-Diffusion/Substrat Diode
- N-Wanne/Substrat Diode
- P-Diffusion/N-Wanne Diode

### 3.2 N-Diffusion/Substrat Diode

Während oben ein idealisiertes Beispiel behandelt wurde, werden jetzt Bauteile besprochen, die Teile von Halbleiter-Chips sind, die im ASIC-Labor Heidelberg entwickelt worden sind. Dabei werden die Dotierprofile des  $0,8\mu\text{m}$ -CMOS Prozesses, mit dem die Chips fabriziert wurden, verwendet, die uns von der Firma *Austria Mikro Systeme International AG* übermittelt worden sind.

Ausgangspunkt ist ein hochreiner Silizium-Einkristall, der mit Boratomen p-grunddotiert ist. Dieser wird dann *Wafer* genannt. In der Abbildung 3.6 ist ein Schnitt senkrecht zur Oberfläche bis ca.  $12\mu\text{m}$  in die Tiefe des Wafers gemacht, und die Akzeptorkonzentration in Abhängigkeit der Tiefe angezeigt. Der Siliziumwafer ist ca.  $250\mu\text{m}$  dick. In der Abbildung ist zu sehen, daß dieser nahe der Oberfläche relativ schwach ( $\approx 3.2 \cdot 10^{14}/\text{cm}^3$ ) mit Bor-Atomen p-dotiert ist. Geht man weiter in die Tiefe, so steigt die Dotierung bis zu ca.  $1.5\mu\text{m}$  an und verbleibt dann bis ca.  $4.5\mu\text{m}$  bei ca.  $7 \cdot 10^{14}/\text{cm}^3$ . Danach steigt die Borkonzentration nochmals an, bis in ca.  $9\mu\text{m}$  Tiefe eine Dotierkonzentration von ca.  $3.4 \cdot 10^{19}/\text{cm}^3$  erreicht ist, die dann auch auf den restlichen  $240\mu\text{m}$  beibehalten wird. Die üblichen Strukturen reichen nicht tiefer als einige bis maximal zehn  $\mu\text{m}$  in den Wafer hinein, so daß diese in einer relativ schwach dotierten Umgebung eingebettet sind. Spricht man vom Wafer in Zusammenhang mit elektrischen Kontakten, so wird dieser oft als *Substrat* bezeichnet. Das Substrat dient als Bezugspunkt der externen Spannungen und liegt meist auf dem niedrigsten elektrischen Potential der ganzen Schaltung.

Läßt man nun von der Oberfläche her Phosphoratome geeigneter Konzentration in den Wafer eindiffundieren, so hat man eine relativ einfache Diode erzeugt. Die Abbildung 3.7 zeigt im



**Abbildung 3.6:** Dotierprofil des Wafers nahe der Oberfläche

linken Teil das Layout einer solchen Diode. Für die Breite der n-Diffusion wurden hier  $15\mu\text{m}$  gewählt. Der Chip-Hersteller läßt diese Größe im allgemeinen offen und der Designer kann sie je nach seinen Erfordernissen wählen, während die Tiefe der Diffusion, die „Aufweichung“ der Grenzschicht und die Höhe der Dotierung vom Herstellungsprozeß bestimmt werden. Diese Größen sind dem Designer zwar bekannt, er hat aber darauf keinen Einfluß.

Im linken Bild der Abbildung 3.7 sieht man die Geometrie der N-Diffusion/Substrat Diode. Zu beachten sind die unterschiedlichen Skalen der Längenangaben<sup>3</sup>. Die Farbskala zeigt die Netto-Dotierung an. Dabei wird nicht zwischen Akzeptor- und Donatorkonzentration unterschieden, das heißt es ist nicht zu erkennen, welche Dotierungsart angezeigt wird, sondern nur der Betrag der Differenz der beiden  $N_{net} = |N_A - N_D|$ . Die beiden „Säulen“ (im Bild violett dargestellt) links und rechts der n-Diffusion dienen zur Ankontaktierung des Substrates, das in diesem Falle gleichzeitig als Anode dient. Um einen niederohmigen Kontakt mit dem Substrat zu erhalten, wird eine zusätzliche hochdotierte p-Diffusion (ca.  $1.4 \cdot 10^{20}/\text{cm}^3$ ) direkt unterhalb der Aluminiumkontakte erzeugt. Die n-Diffusion wird ebenfalls mit einem Aluminiumkontakt versehen, der als Kathode bezeichnet wird. Im rechten Bild der Abbildung 3.7 ist ein Schnitt, der im linken als dünner Strich in der Mitte der Kathode gekennzeichnet ist, durch die Mitte der n-Diffusion gezeigt. Die rote Kurve stellt die Donatorkonzentration dar und die grüne die Akzeptorkonzentration. Die blaue Kurve entspricht der Farbskala im linken Bild und ist die Überlagerung der beiden anderen. Hier sieht man, daß im oberen Teil die Netto-Dotierung fast ausschließlich von den Phosphoratomen bestimmt wird, obwohl dort auch die Bor-Dotierung des Wafers vorhanden ist, allerdings mit einem Verhältnis von  $N_D/N_A \approx 10^6$  – in Worten: auf eine Million Phosphoratomme kommt ein Boratom. Man kann sie also an dieser Stelle als nicht existent betrachten, da sie durch die Überzahl der freien Elektronen, aufgrund der hohen Donatorkonzentration, vollständig kompensiert sind. Tiefer

<sup>3</sup>Um einen Eindruck für die wahren Größenverhältnisse zu bekommen, sind im Anhang A maßstabsgetreue Bilder aller in der vorliegenden Arbeit besprochenen Bauelemente aufgeführt.

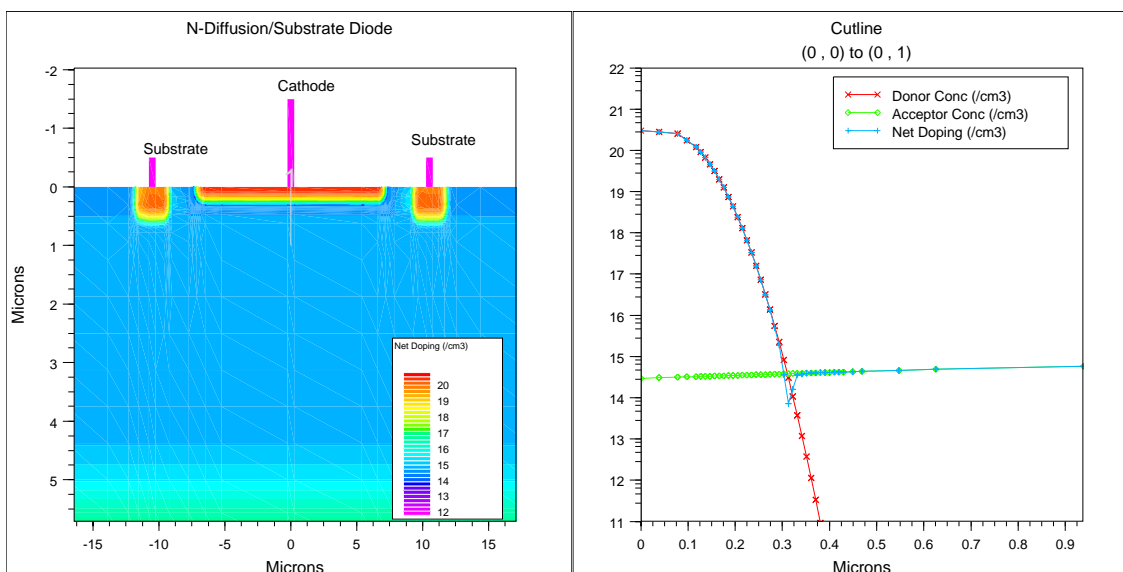


Abbildung 3.7: Dotierprofil der N-Diffusion/Substrat Diode

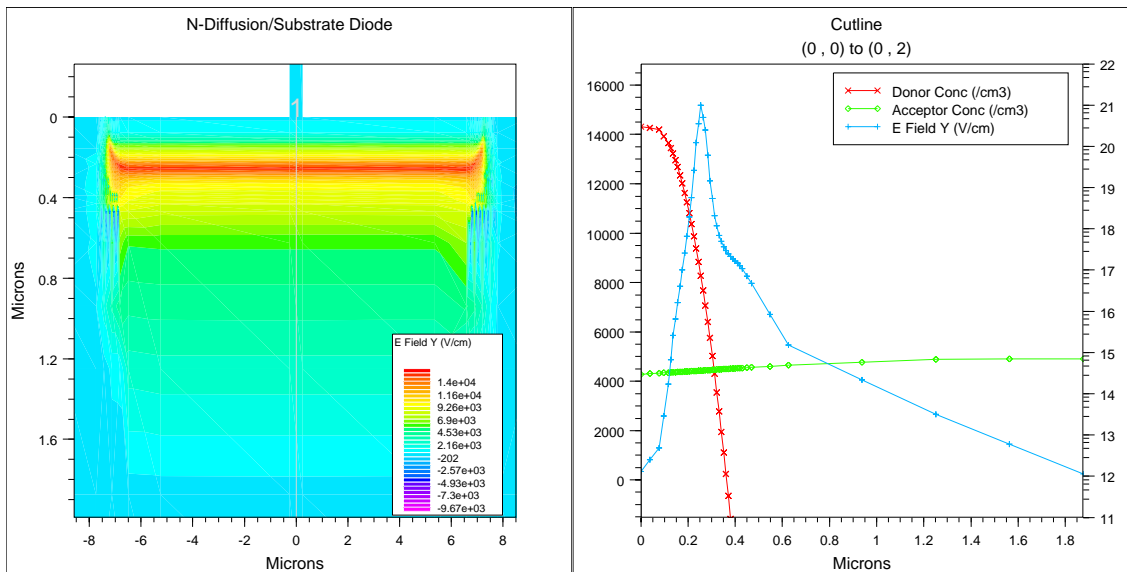
im Substrat (ca.  $0.35\mu m$ ) überwiegt dann die Akzeptordichte die der Donatoren, und die Netto-Dotierung wird durch die p-Dotierung des Substrates bestimmt.

Betrachtet man nun die Raumladungszone, so wird diese nicht mehr symmetrisch beiderseits des pn-Überganges verlaufen. Aufgrund der sehr hohen Donatorkonzentration im n-Gebiet rekombinieren die über den pn-Übergang diffundierten Löcher sehr schnell mit den Elektronen im n-Gebiet, während die in die p-Zone (hier also das Substrat) diffundierten Elektronen eine relativ schwache p-Dotierung „sehen“ und somit nicht so viele Rekombinationsmöglichkeiten haben und dadurch weiter von der Übergangsschicht wegdiffundieren. Die Raumladungszone wird sich daher vor allem in den schwächer dotierten Bereich erstrecken.

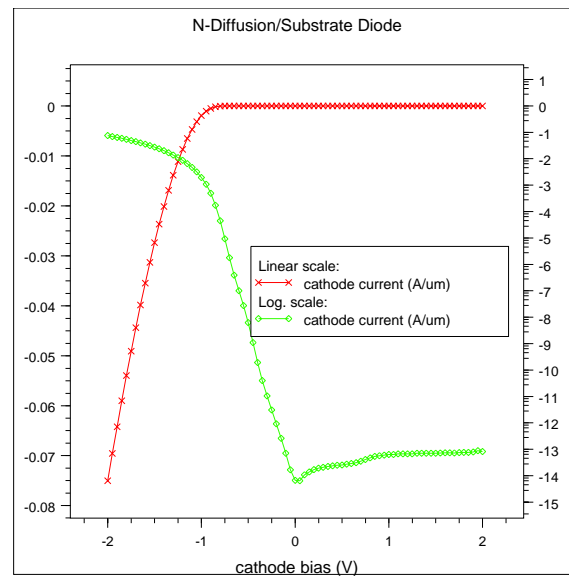
Die Abbildung 3.8 verdeutlicht diesen Sachverhalt. Der linke Teil der Abbildung zeigt einen Ausschnitt der N-Diffusion/Substrat Diode. Die Farbskala stellt die y-Komponente des elektrischen Feldes in der Nähe der Übergangszone dar. Es ist deutlich zu sehen, wie sich das Feld sehr viel mehr in die schwächer dotierte p-Zone erstreckt (ca.  $1.5\mu m$ ), während es in der hoch dotierten n-Zone relativ schnell (ca.  $0.2\mu m$ ) verschwindet. Der rechte Teil der Abbildung zeigt das Gleiche in einem Schnitt durch die Diode an. Die elektrische Feldstärke (blau) verläuft im n-Gebiet ( $0 < y < 0.25\mu m$ ) sehr viel steiler als im p-Gebiet, was bedeutet, daß die Raumladungen im n-Gebiet räumlich dichter liegen als im p-Gebiet. Zur Übersicht sind noch die Donatorkonzentration (rot) und die Akzeptorkonzentration (grün) mit aufgetragen. Die pn-Übergangszone befindet sich ungefähr an der Stelle, an der  $N_D = N_A$  wird. Dort erreicht auch das elektrische Feld seinen maximalen Wert. Zuletzt ist in der Abbildung 3.9 die Kennlinie der Diode zu sehen. Da man auf dem Chip über das Potential der Anode nicht verfügen kann<sup>4</sup>, muß die Kathodenspannung variiert werden, um die Kennlinie zu erhalten.

Der Vorteil dieser Diode ist ihr geringer Aufwand in der Herstellung und ihr kleiner Platzbe-

<sup>4</sup>In der Simulation wäre das ohne weiteres möglich, da hier nur einzelne Bauteile simuliert werden. Es wurde aber mit Rücksicht auf die realen Gegebenheiten das Substratpotential auf 0V gelegt.



**Abbildung 3.8:** y-Komponente des elektrischen Feldes der N-Diffusion/Substrat Diode im thermodynamischen Gleichgewicht



**Abbildung 3.9:** Kennlinie der N-Diffusion/Substrat Diode

darf. Der Nachteil besteht darin, daß als Anode das Substrat dient. Dadurch ist das Bauteil relativ schlecht von der Umgebung abgeschirmt und etwaige diffundierende Ladungsträger von benachbarten Strukturen können das elektrische Verhalten der Diode stören. Außerdem muß die Anode, die in diesem Falle einfach das Substrat ist, den Einschränkungen des Substrates gehorchen. Das heißt die Anode muß auf dem niedrigsten Potential der Schaltung liegen, wodurch die Diode nur in Sperrichtung betrieben werden kann.

Beim nächsten Beispiel handelt es sich um einen ähnlichen Aufbau wie im vorhergehenden, mit dem Unterschied, daß als n-Gebiet nun nicht die relativ dünne n-Diffusion dient, sondern ein n-Gebiet, das zum einen nicht so stark dotiert ist, und zum anderen tiefer in das Substrat reicht.

### 3.3 N-Wanne/Substrat Diode

Ausgangspunkt ist wieder der Silizium-Waferrohling aus Abbildung 3.6. In diesen werden nun Donatoren (hier Phosphoratome) eindiffundiert, so daß sie ein n-Gebiet bilden, das sich bis in ca.  $4\mu\text{m}$  Tiefe erstreckt. Dieses Gebiet wird als *N-Wanne* (engl.: n-well) bezeichnet. Der Name leitet sich daraus ab, weil die Struktur für andere Bauteile dazu dient, eine n-Umgebung zu erzeugen, in die dann weitere Diffusionen eingebracht werden<sup>5</sup>. Diese Diffusionen sind dann wie in einer „Wanne“ eingebettet. In der Abbildung 3.10 ist das Layout der N-Wanne/Substrat Diode zu sehen. Die Farbskala im linken Bild stellt die Netto-Dotierung

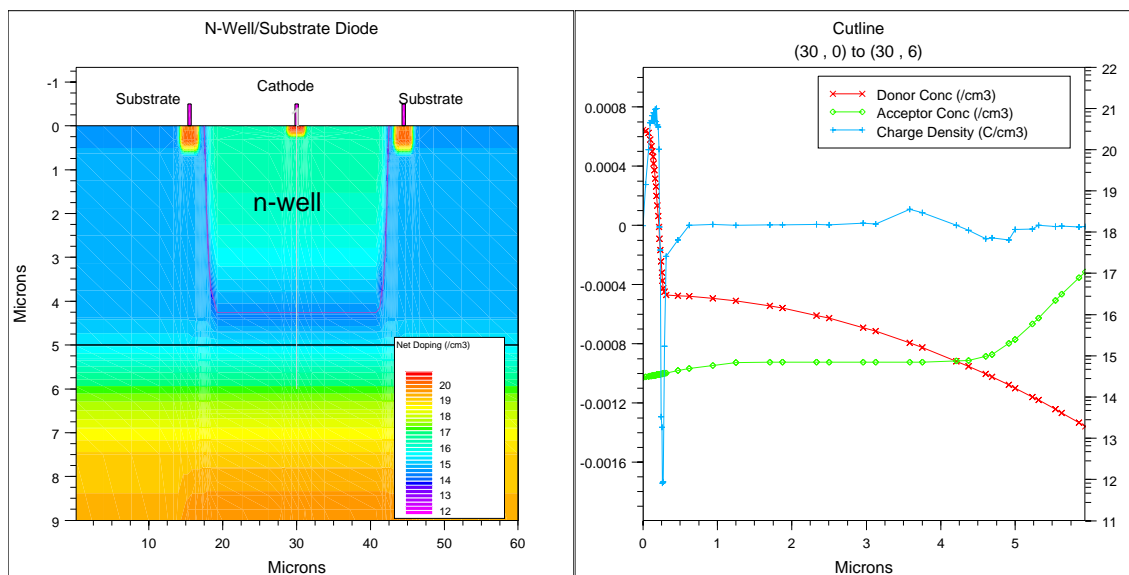
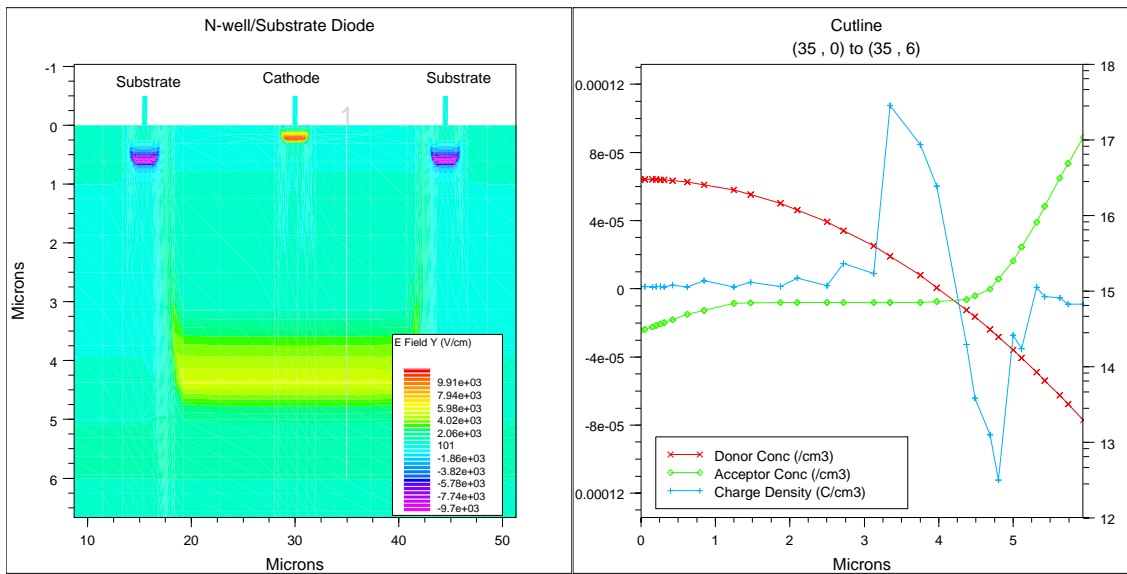


Abbildung 3.10: Dotierprofil und Raumladung der N-Wanne/Substrat Diode

dar. Der rote „Fleck“ unterhalb des Kathodenkontaktes ist eine hochdotierte n-Zone, um die Wanne niederohmig zu kontaktieren. Analog dazu dienen die beiden hochdotierten p-Zonen rechts und links der Wanne unterhalb des jeweiligen Substratkontaktes dazu, das Substrat niederohmig mit dem entsprechenden Aluminiumkontakt zu verbinden. Wie man an der Farbskala erkennt, ist die Wanne relativ schwach mit ca.  $3 \cdot 10^{16} / \text{cm}^3$  dotiert. Als n-Wanne ist sie mit Donatoren, hier speziell mit Phosphor, dotiert. Im rechten Bild ist ein Schnitt mitten durch die Wanne gemacht. Die rote Kurve zeigt die Donatorkonzentration. Direkt unter der Oberfläche an dieser Stelle ( $x = 30\mu\text{m}$ ) wird sie durch die hochdotierte Zone der Ankontaktierung bestimmt und fällt bei  $y \approx 0.3\mu\text{m}$  auf die Wannendotierung ab. Grün eingezeichnet ist die Akzeptorkonzentration, die diejenige des Substrates ist und schon von der

<sup>5</sup>Später wird die Wanne beim vertikalen pnp-Bipolar Transistor als Basis desselben benutzt

N-Diffusion/Substrat Diode her bekannt ist. Bemerkenswert ist der Verlauf der Raumladungsdichte, der blau gekennzeichnet ist. Es ist an der Stelle  $x = 30\mu\text{m}$  im Bereich  $0 < y < 0.3\mu\text{m}$  eine Raumladungszone zu erkennen, obwohl dort kein pn-Übergang vorhanden ist. Es herrscht dort aber ein Konzentrationsgradient bezüglich der Elektronenkonzentration (an der Donorkonzentration erkennt man den Übergang von  $\approx 10^{20}/\text{cm}^3$  Donatoren auf  $3 \cdot 10^{16}/\text{cm}^3$ ), so daß auch hier die Elektronen durch Diffusion versuchen, diesen abzubauen. Die Diffusion bewirkt wiederum eine Ladungsanhäufung jenseits des Überganges, und es entsteht ein elektrisches Feld, das der Diffusion entgegenwirkt. Erst an der Stelle  $y \approx 4\mu\text{m}$  befindet sich der „normale“ pn-Übergang, ebenfalls mit einer Raumladungszone, deren Maximum und Minimum relativ klein, deren räumliche Ausdehnung aber aufgrund der schwachen Dotierung größer ist. Die y-Komponente des elektrischen Feldes ist im linken Teil der Abbildung 3.11



**Abbildung 3.11:** y-Komponente des elektrischen Feldes und Raumladungen der N-Wanne/Substrat Diode am pn-Übergang

gezeigt. Dort kann man das relativ starke, aber räumlich scharf begrenzte Feld als rote Farbmarkierung unterhalb des Kathodenanschlusses erkennen. Ebenso sieht man, daß Analoges für die hochdotierten Ankontaktierungen des Substrates unterhalb der Substratkontakte gilt. Die Richtung des Feldes ist zwar genau entgegengesetzt, der Betrag aber ungefähr gleich.

Um die Raumladungszone am pn-Übergang zu verdeutlichen, wurde nochmals ein Schnitt durch die Wanne definiert, wobei nun aber an der Stelle  $x = 35\mu\text{m}$  das Tiefenprofil gezeigt wird. Die hochdotierte Ankontaktierung ist nun nicht mehr zu sehen, und ebenso ist auch die Raumladung im oberen Bereich verschwunden. Im Bereich von ca.  $4.3\mu\text{m}$  sieht man jetzt deutlicher die eigentliche pn-Übergangsschicht mit ihren Raumladungen. Hier hat sich im Vergleich zum obigen Bild zur Verdeutlichung lediglich die Skala geändert, wobei Größe und räumliche Ausdehnung gleich geblieben sind. Beim Betrieb dieser Diode auf dem Chip gelten die gleichen Einschränkungen, wie für die N-Diffusion/Substrat Diode, denn auch hier ist die Anode das Substrat. Das Potential des Substrates ist fest und kann nicht variiert werden.

Will man sich aber als Designer die Freiheit in der Wahl des Potentials für die Anode erhalten

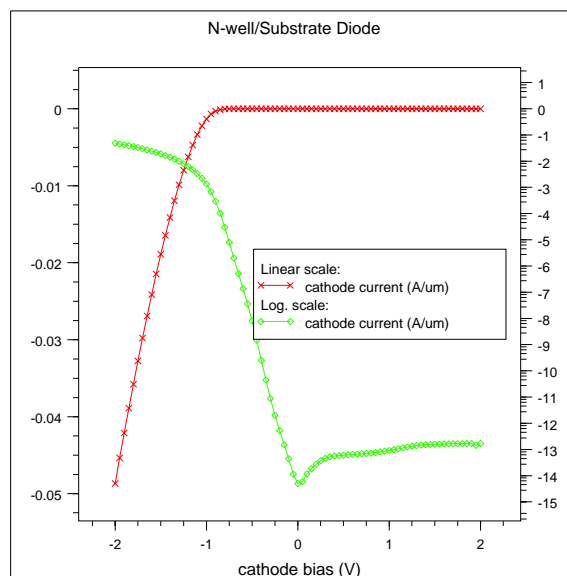


Abbildung 3.12: Kennlinie der N-Wanne/Substrat Diode

(das heißt insbesondere, eine Diode in Vorwärtsrichtung betreiben), so muß man zu einer etwas anderen Struktur übergehen. Im folgenden wird eine sogenannte *P-Diffusion/N-Wanne Diode* vorgestellt, die den obigen Einschränkungen für die Anode nicht mehr unterliegt.

### 3.4 P-Diffusion/N-Wanne Diode

Um eine P-Diffusion/N-Wanne Diode zu erzeugen, geht man vom Silizium-Wafer mit eingebrachter N-Wanne aus. Der Unterschied besteht bis hierher lediglich darin, daß die Ankontaktierung nun nicht mehr in der Mitte der Wanne erfolgt, sondern jeweils links und rechts am Rand der Wanne eine hochdotierte n-Zone in die Wanne eindiffundiert ist. Diese beiden Ankontaktierungen bilden gemeinsam die Kathode dieser Diode. In einem weiteren Prozeßschritt werden in die N-Wanne Akzeptoren (hier Bor) eingebracht, die dann die Anode der Diode bilden.

In der Abbildung 3.13 ist im linken Teil das Layout der P-Diffusion/N-Wannen Diode gezeigt. Die Farbskala stellt die Netto-Dotierung ( $N_{net} = |N_D - N_A|$ ) dar. Die N-Wanne ist links und rechts mit hochdotierten n-Diffusionen versehen, die einen niederohmigen Kontakt zur Wanne sicherstellen. Rechts und links neben der Wanne sind hochdotierte p-Gebiete in das Substrat eingebracht, um dieses zu kontaktieren. Im rechten Teil der Abbildung ist ein Schnitt mitten durch die Diode gezeigt. Hier kann man die Wannen-Dotierung (rote Kurve) an der rechten, logarithmischen Skala ablesen. Sie beträgt, wie oben erwähnt, maximal  $3 \cdot 10^{16}$  Donatoren/cm<sup>3</sup> und fällt relativ flach ab. Die p-Dotierung der Anode (grüne Kurve) ähnelt dem Dotierprofil des n-Gebietes der N-Diffusion/Substrat Diode sehr stark, mit dem Unterschied, daß sie natürlich aus Akzeptoren besteht. Bei ca.  $0.8\mu\text{m}$  geht die Bordotierung in die Grunddotierung des Wafers über. Die blaue Kurve zeigt letztlich die Raumladungsdichte ( $\rho = q(N_D + p - N_A - n)$ ), wobei man wieder erkennen kann, daß diese sich vorwiegend in

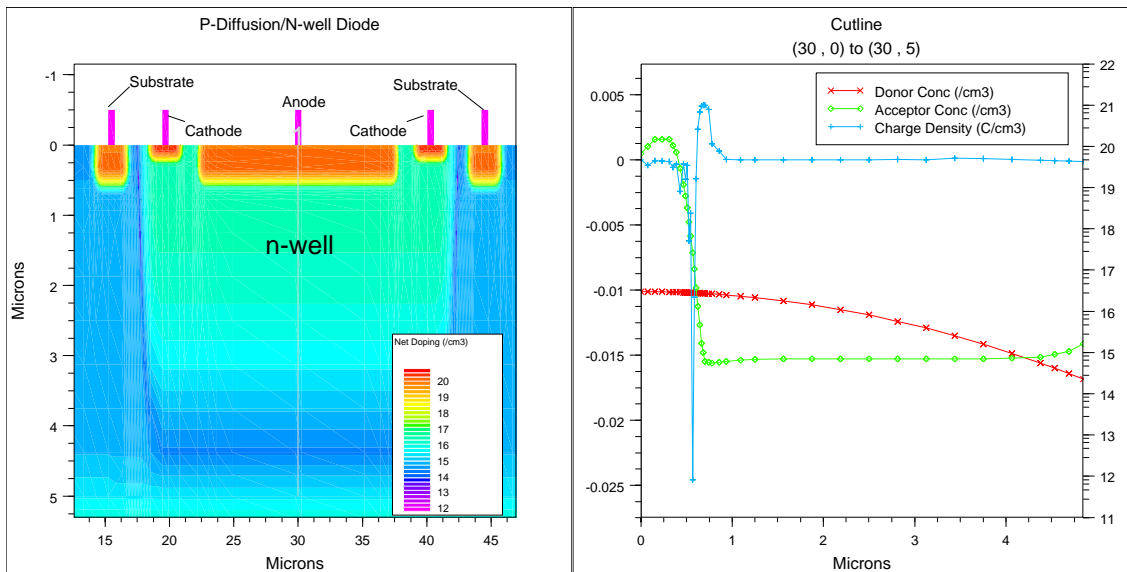


Abbildung 3.13: Layout und Dotierprofil der P-Diffusion/N-Wannen Diode

die niedriger dotierte Zone erstreckt.

Wie oben erwähnt hat diese Diode den Vorteil, daß Anode und Kathode vom Substrat getrennt sind, und somit mit „beliebigen“ Spannungen versehen werden können. Dabei gilt natürlich immer noch, daß das Substrat meist das niedrigste Potential hat. Das ist aber für den Betrieb der Diode keine Einschränkung mehr, denn die Anodenspannung der p-Diffusion ist nun für jeden gewünschten Betrieb der Diode frei wählbar, das heißt, die Anodenspannung kann größer oder kleiner sein als die Kathodenspannung (also das externe Potential der Wanne). Der offensichtliche Nachteil dieser Diode ist ihr komplizierterer Herstellungsprozeß und der größere Platzbedarf gegenüber der N-Diffusion/Substrat Diode.



# Kapitel 4

## Bipolar Transistoren

Im vorhergehenden Kapitel sind Bauteile besprochen worden, die genau einen pn-Übergang besitzen. Sie werden als Dioden bezeichnet und ihre unsymmetrische Strom-Spannungscharakteristik bezüglich Durchlaß- und Sperrspannung wird gerade durch diesen pn-Übergang bestimmt. Da es bei der Diode nur zwei verschiedene Gebiete gibt, hat man auch nur zwei elektrische Anschlüsse zur Verfügung, um mittels externer Felder das Verhalten des Bauteiles zu beeinflussen.

Fügt man dem Bauteil einen weiteren pn-Übergang zu, so erhält man einen Transistor. Dabei hat man prinzipiell zwei Möglichkeiten. Entweder man wählt die Schichtfolge „npn“ oder „pnp“. Im ersten Fall erhält man einen npn-Transistor, entsprechend im zweiten einen pnp-Transistor.

Wenn man sich die P-Diffusion/N-Wanne Diode genauer anschaut, so kann man erkennen, daß hier schon eine pnp-Schichtfolge realisiert ist:

- p: P-Diffusion
- n: N-Wanne
- p: p-Dotierung des Wafers.

Die Abbildung 4.1 zeigt das Schaltbild eines pnp-Bipolar Transistors. Da man nun drei verschiedene Gebiete, entsprechend der Schichtfolge hat, sind zur elektrischen Spannungszuführung ebenfalls drei Kontakte notwendig. Diese werden als Emitter (E), Kollektor (C) und Basis (B) bezeichnet.

Da der Herstellungsprozeß, nach dessen Dotierprofile die bisher besprochenen Strukturen erzeugt wurden, ein CMOS-Prozeß ist, also im wesentlichen auf die Herstellung von MOS-Strukturen abzielt, wird der durch die obige Schichtfolge entstandene Bipolar-Transistor als *parasitärer* pnp-Transistor bezeichnet. Weiterhin bewegen sich die Ladungsträger beim Betrieb dieses Transistors senkrecht zur Oberfläche des Chips in das Substrat hinein, weshalb er zusätzlich als *vertikaler* pnp-Transistor bezeichnet wird. Man erhält also letztlich einen *parasitären vertikalen pnp-Transistor*. Im nächsten Abschnitt wird dieser Transistor genauer besprochen.

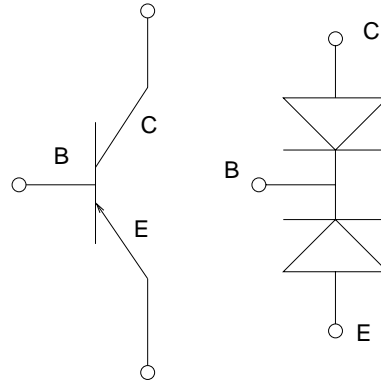


Abbildung 4.1: Schaltungszeichen des pnp-Bipolar Transistors und Dioden Ersatzschaltbild

### 4.1 Der Parasitäre Vertikale PNP-Transistor

Wie oben schon erwähnt werden zur Herstellung des parasitären vertikalen Bipolar Transistors (im folgenden als *pnp-Transistor* bezeichnet) keine neuen Strukturen gebraucht. Im linken Teil der Abbildung 4.2 ist in farblicher Kennzeichnung die Netto-Dotierung des pnp-Transistors gezeigt. Vergleicht man dazu die Abbildung 3.13, so kann man keine Unterschiede in der Dotierung feststellen. Der einzige Unterschied besteht in der Benennung der Kontakte, die natürlich willkürlich ist und sich lediglich nach den üblichen Bezeichnungen richtet. Im

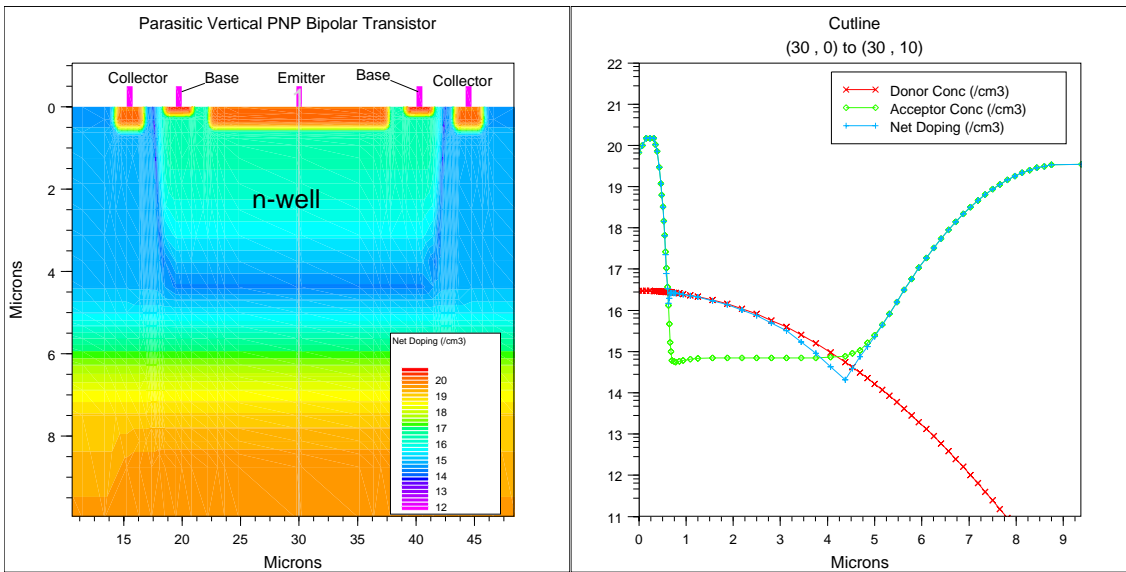
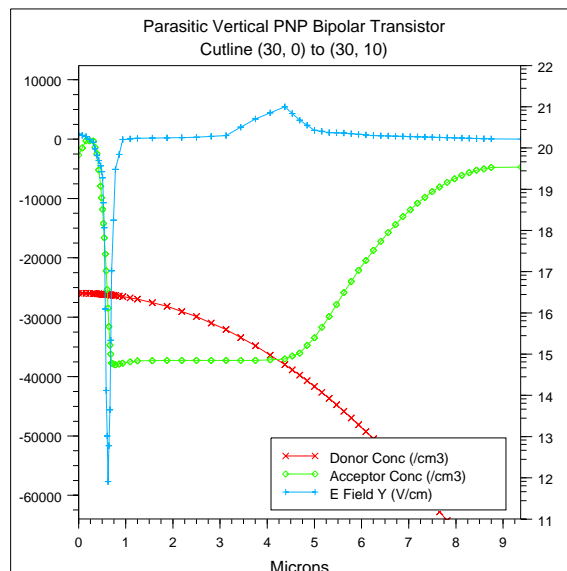


Abbildung 4.2: Layout des parasitären vertikalen pnp-Transistors

Vergleich sieht man, daß die hochdotierte p-Diffusion, die bei der P-Diffusion/N-Wanne Diode als Anode diente, hier den Emitter darstellt. Weiterhin wird die N-Wanne, die bei der Diode die Kathode war nun zur Basis des pnp-Transistors. Und letztlich dient das Substrat jetzt als Kollektor. Im rechten Teil der Abbildung sind die Konzentrationen der Donatoren (rot)

und Akzeptoren (grün) längs eines Schnittes in der Mitte des pnp-Transistors aufgetragen. Die blaue Kurve zeigt die Netto-Dotierung ( $N_{net} = |N_D - N_A|$ ), die der Farbskala im linken Teil der Abbildung entspricht. Auch hier gibt es keinen Unterschied zu dem Schnitt durch die P-Diffusion/N-Wanne Diode. Das bedeutet also für den Transistor, daß der Emitter am höchsten dotiert ist, während Basis und Kollektor schwach dotiert sind. Das hat Konsequenzen für die Feldverhältnisse, die an den beiden pn-Übergängen herrschen. Die Abbildung 4.3 zeigt den gleichen Schnitt wie oben, aber es ist die y-Komponente des Feldes (blaue Kurve) im thermodynamischen Gleichgewicht (alle externen Spannungen  $0V$ ) aufgetragen. Man kann

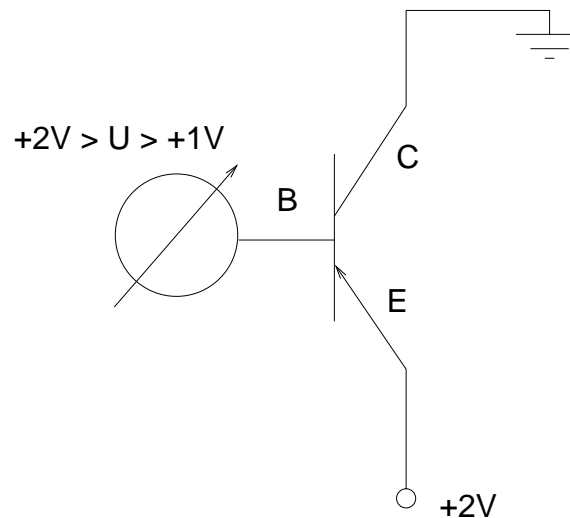


**Abbildung 4.3:** y-Komponente des elektrischen Feldes an den beiden pn-Übergängen

die relativ hohe elektrische Feldstärke am pn-Übergang der „Emitter-Basis Diode“ erkennen und ebenfalls das schwache Feld, aufgrund der schwachen Dotierung von Basis und Kollektor, im Bereich der „Basis-Kollektor Diode“.

Solange die Basis (N-Wanne) und der Kollektor (Substrat) das gleiche externe Potential haben, fungiert das Bauteil als Diode, wie im vorherigen Kapitel beschrieben. Erst wenn sich das externe Potential der Basis von dem des Kollektors unterscheidet, geht man zum Transistorbetrieb über. Weil der Kollektor vom Substrat gebildet wird, gilt für diesen die Bedingung, die immer für das Substrat-Potential erfüllt sein muß, daß er auf dem niedrigsten Potential der Schaltung liegt. Im folgenden wird daher das Kollektorpotential immer  $0V$  betragen.

Um die Wirkungsweise des Transistors zu verdeutlichen, wird die sogenannte *Common Emitter* Schaltung betrachtet. Sie wird deshalb so genannt, weil das gemeinsame Bezugspotential dasjenige des Emitters ist. Die Abbildung 4.4 zeigt die Common Emitter Schaltung und die externen Potentiale. Zunächst wird dem Basisanschluß eine Spannung von  $2V$  auferlegt. Dadurch sind sowohl die Basis-Kollektor Diode als auch die Basis-Emitter Diode in Sperrichtung geschaltet, und es fließen über die Basis die Rückwärts-Sättigungsströme der beiden Dioden ab. Nun wird der Emitter auf  $+2V$  gelegt, wodurch zwischen Emitter und Basis keine Potentialdifferenz mehr herrscht. Das entspricht nun der Konfiguration der N-



**Abbildung 4.4:** Externe Spannungen der Common Emitter Schaltung und Dioden-Ersatzschaltbild

Wanne/Substrat Diode mit  $2V$  Rückwärtsspannung. Verringert man nun die Basisspannung, so wird die Basis-Emitter Diode in Vorwärtsrichtung gepolt, wodurch die Löcherkonzentration in der Basis zunimmt. Da die Basis n-dotiert ist, sind die Löcher dort Minoritäten. Wäre nun die Basis hoch dotiert, würden die Minoritäten (also die vom Emitter in die Basis injizierten Löcher) schnell mit den „freien“ Elektronen der Basis rekombinieren, über den Basiskontakt abfließen und man hätte nichts weiter als die P-Diffusion/N-Wanne Diodenkonfiguration in Vorwärtsrichtung. Für das Zustandekommen der Transistorfunktion sind nun zwei Faktoren wesentlich:

1. Wie in der Abbildung 4.3 zu erkennen ist, herrscht in dem Gebiet der Basis ein feldfreier Raum, was bedeutet, daß die vom Emitter in die Basis injizierten Löcher keinem Potentialgefälle unterliegen und so kräftefrei durch die Basis diffundieren können.
2. Die Basis darf nicht zu hoch dotiert sein, da sie sonst den Löchern zu viele Rekombinationszentren böte, wodurch diese „vernichtet“ würden. Außerdem darf die Basis nicht zu breit sein, da es dann auch zu viele Möglichkeiten für die Löcher gäbe, zu rekombinieren.

Wenn nun diese Bedingungen gegeben sind, dann gelangen die Löcher in die Nähe der Basis-Kollektor Raumladungszone, in der das elektrische Feld vom n- zum p-Gebiet gerichtet ist (siehe Abb. 4.3). Durch dieses Feld werden die Löcher in die Kollektorzone gezogen und von diesem, aufgrund seines negativsten Potentials, über den Kontakt abfließen. Die Abbildung 4.5 zeigt den PNP-Transistor in drei verschiedenen Spannungszuständen. Die dünnen violetten Linien zeigen die pn-Übergangszonen an. Im oberen Bild liegen am Emitter und an der Basis jeweils  $+2V$  an, während der Kollektor, wie in allen drei Bildern, auf Null Volt liegt. Das entspricht wie oben erwähnt der N-Wanne/Substrat Diodenkonfiguration mit  $2V$  Sperrspannung, weshalb sich die Löcherkonzentration in der Nähe der Basis-Kollektor Diode etwas erniedrigt hat. Im mittleren Bild beträgt die Emitter-Spannung ebenfalls  $+2V$ , die Basis-Spannung aber nur noch  $+1.5V$ , was bedeutet, daß die Basis-Emitter Diode des

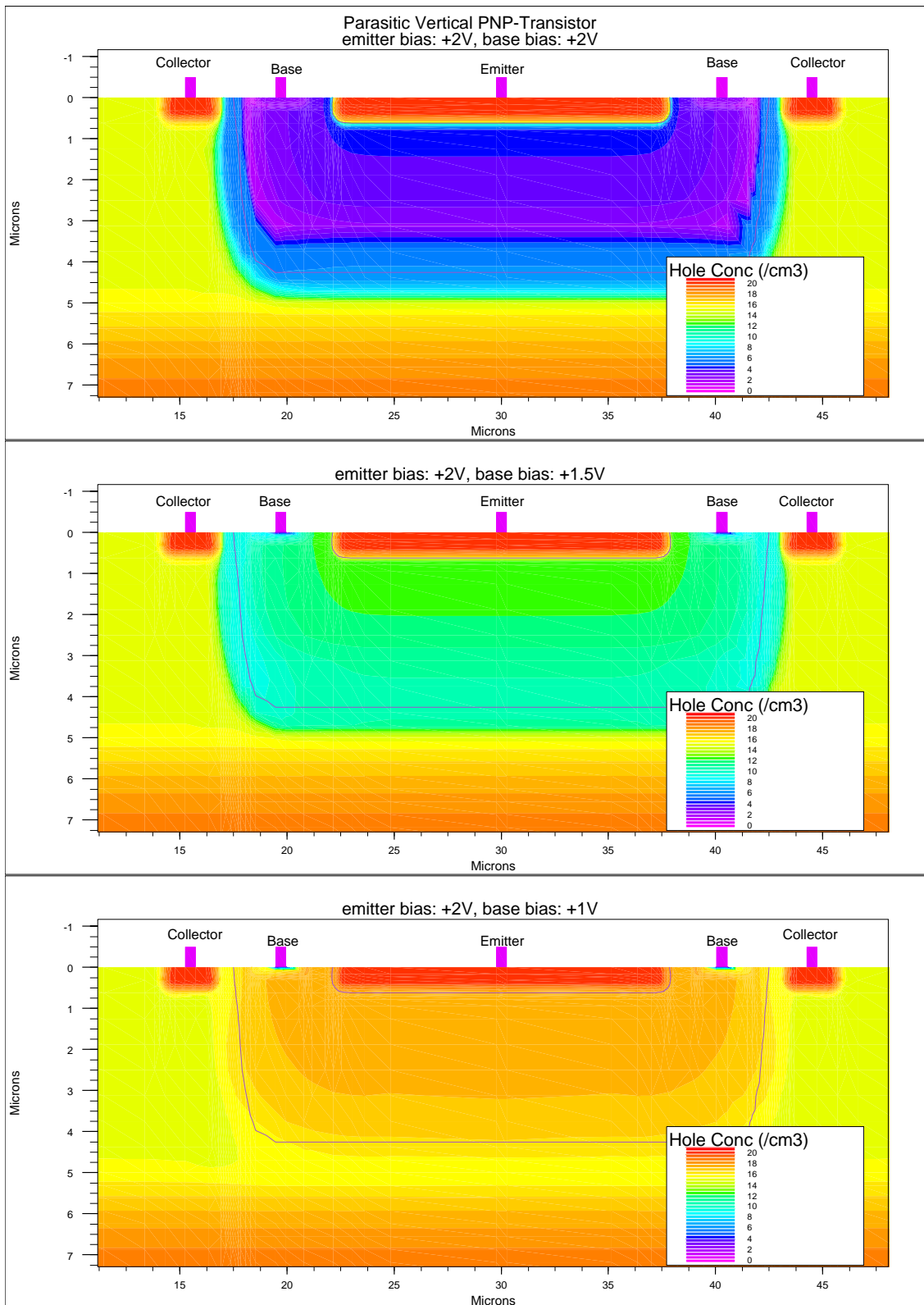
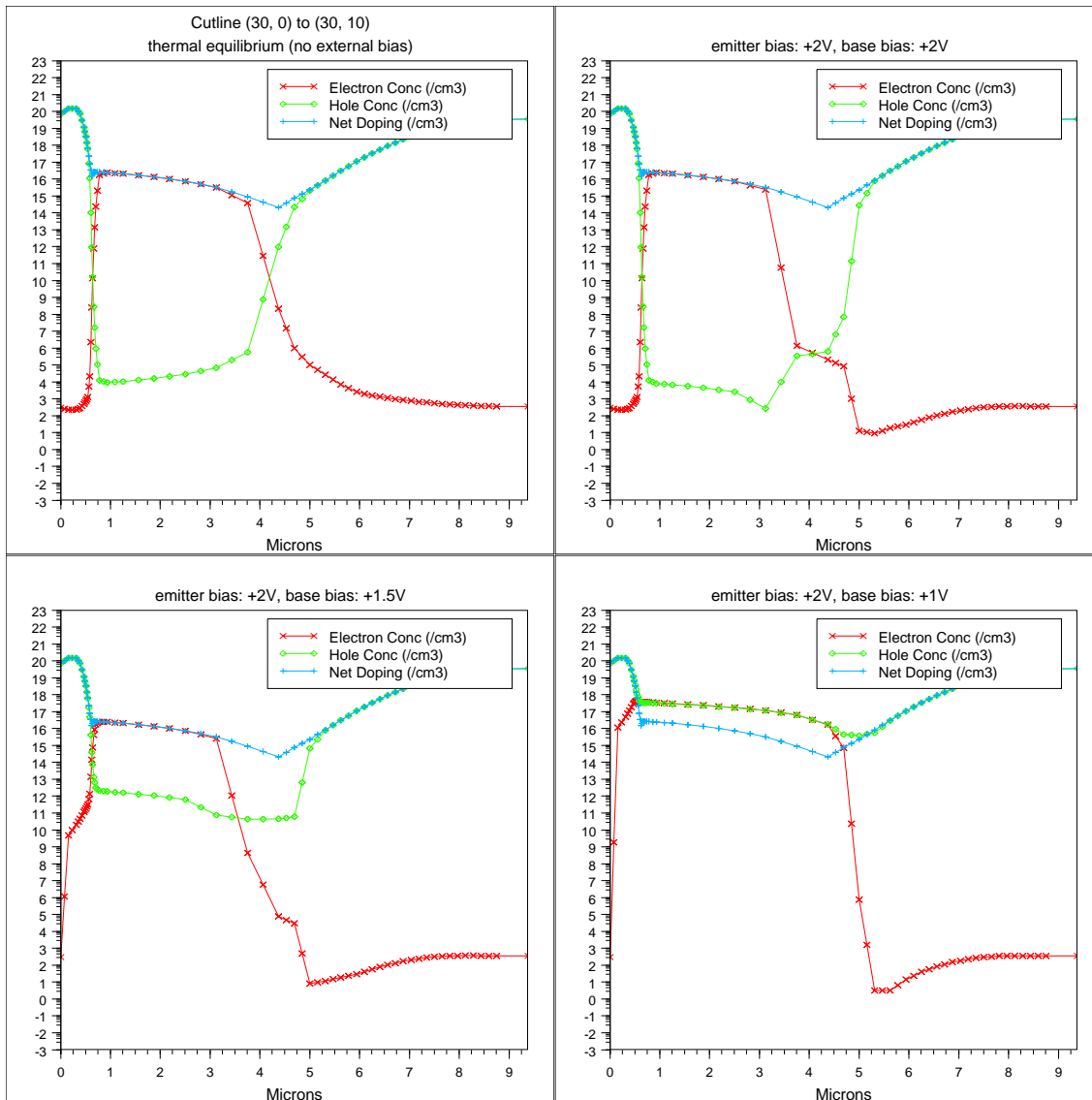


Abbildung 4.5: Löcherkonzentration des parasitären vertikalen PNP-Transistors bei verschiedenen externen Spannungen

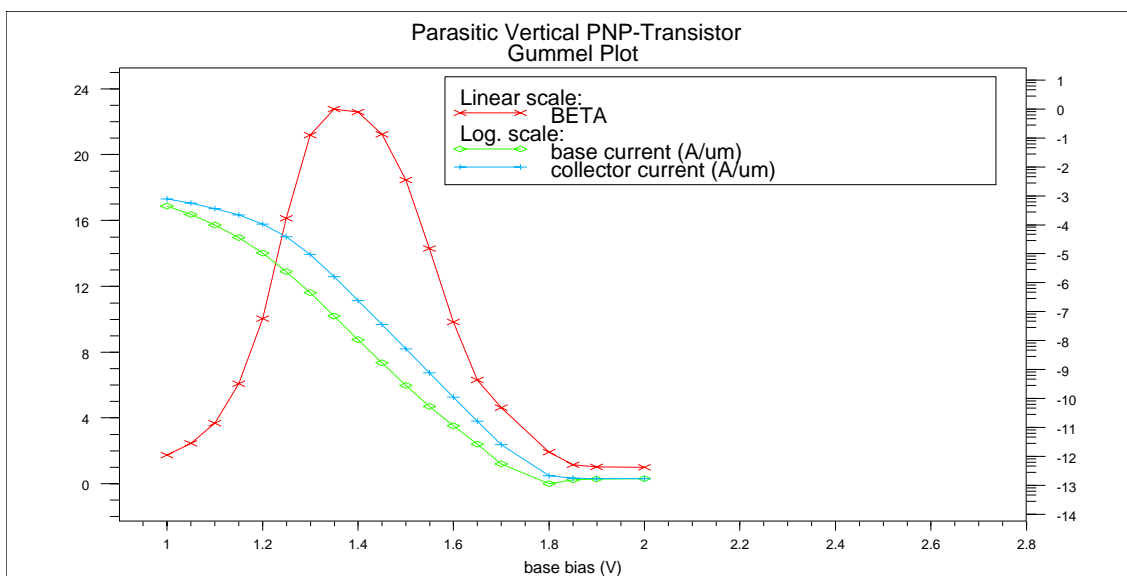
PNP-Transistors nun mit einer Durchlaßspannung von  $0.5V$  betrieben wird. Das macht sich deutlich in der Löcherkonzentration in der Basis, die dort Minoritäten sind, bemerkbar. Verringert man die externe Spannung an der Basis noch weiter, so kommt man zum unteren Bild, in dem die Basisspannung nur noch  $+1V$  beträgt, und somit die Basis-Emitter Diode mit einer Durchlaßspannung von  $1V$  geschaltet ist. Die Löcherkonzentration in der Basis steigt dramatisch an, und der Transistor ist nun voll angesteuert.



**Abbildung 4.6:** Löcher- und Elektronenkonzentration des parasitären vertikalen PNP-Transistors bei verschiedenen externen Spannungen

Um einen besseren Einblick in die Ladungsträgerkonzentrationen zu erhalten, wurde mitten durch den Transistor ein Schnitt gelegt und die Löcher- und Elektronenkonzentration längs der Schnittlinie in der Abbildung 4.6 aufgetragen. Um die pn-Übergänge zu markieren, wurde zusätzlich die Netto-Dotierung mit eingetragen, die an ihren „Knickstellen“ die Übergänge

vom p- zum n-Gebiet bzw. vom n- zum p-Gebiet kennzeichnet. Das linke obere Bild zeigt die Ladungsträgerverteilung im thermodynamischen Gleichgewicht, also ohne äußere Spannungen. Zunächst sei die Elektronenkonzentration (rote Kurve) betrachtet. Sie ist entsprechend der Dotierung im Emitter (ca.  $10^{20}/\text{cm}^3$  Akzeptoren) sehr gering und steigt im Gebiet der Basis auf ca.  $3 \cdot 10^{16}$  an. Danach fällt sie wieder im Kollektor auf die Minoritätenkonzentration ab. Im rechten oberen Bild liegen am Emitter und an der Basis jeweils  $+2V$  an, so daß über der Basis-Emitter Diode keine Spannung abfällt, und sich somit an dieser Stelle an den Ladungsträgerkonzentrationen im Vergleich zum vorherigen Bild nichts ändert. Die Basis-Kollektor Diode ist nun aber in Sperrichtung geschaltet, wodurch sich die Raumladungszone an diesem pn-Übergang verbreitert, was sich in der Verringerung der Ladungsträgerdichten ausdrückt. Im linken unteren Bild ist der Basis-Emitter Übergang mit  $0.5V$  Durchlaßspannung geschaltet. Wie in Kapitel 3 beschrieben, unterstützt die Durchlaßspannung das Diffundieren der Ladungsträger in das jeweils andere Gebiet (die Löcher in die n-Zone bzw. die Elektronen in die p-Zone). Dies macht sich durch eine Erhöhung der Löcherkonzentration in der Basis um ca. acht Größenordnungen bemerkbar. Verringert man die Basisspannung noch mehr, das heißt, die Basis-Emitter Diode wird noch weiter in Vorwärtsrichtung betrieben, so kommt man zum Bereich der „Hochinjektion“, den das rechte untere Bild darstellt. Es ist zu erkennen, daß die Löcherkonzentration höher als die Elektronenkonzentration der Basis im thermischen Gleichgewicht ist, weshalb man von Hochinjektion spricht. In der Abbil-



**Abbildung 4.7:** Basisstrom, Kollektorstrom und Stromverstärkung des parasitären vertikalen PNP-Transistors

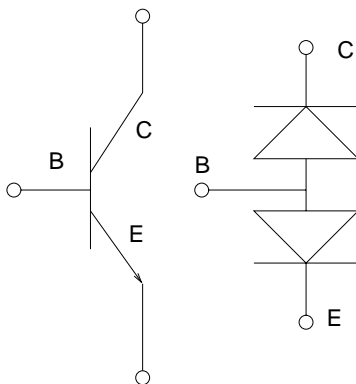
dung 4.7 ist in logarithmischer Skala der Basisstrom (grün) und der Kollektorstrom (blau) in Abhängigkeit der Basisspannung aufgetragen, wobei der Emitter auf  $+2V$  und der Kollektor auf  $0V$  liegt. Dieser sogenannte *Gummel-Plot* beinhaltet also die drei soeben beschriebenen Spannungszustände. Zusätzlich ist die Stromverstärkung  $\beta$  (blau) mit eingezeichnet. Die Stromverstärkung ergibt sich aus dem Verhältnis von Kollektorstrom und Basisstrom zu  $\beta = I_C/I_B$ . Sie beträgt bei diesem Transistor, wie in der Abbildung 4.7 zu erkennen ist, bei ihrem maximalen Wert  $\beta \approx 22$ .

Wie oben angedeutet ist auch dieses Bauelement nur beschränkt einsetzbar. Wie man am Layout des Transistors sehen kann, dient als Kollektor das Substrat, was bedeutet, daß dieser den Einschränkungen des Substrates unterliegt. Das heißt der Kollektor muß erstens das niedrigste Potential der ganzen Schaltung besitzen, und zweitens ist das Potential dadurch festgelegt und kann nicht variiert werden. Der Vorteil ist der, daß mit den Strukturen des CMOS-Prozesses dennoch ein Bipolar-Transistor erzeugt werden kann, und der Designer somit mit ein und demselben Prozeß trotzdem, wenn auch relativ rudimentär, in seinem CMOS-Design Bipolar-Transistoren verwenden kann.

Im nächsten Abschnitt wird ein Bipolar-Transistor vorgestellt, dessen Anschlüsse vom Substrat entkoppelt sind. Dazu muß aber ein anderer Prozeß mit anderen Strukturen, die im CMOS-Prozeß nicht zur Verfügung stehen, herangezogen werden. Weil dieser sowohl CMOS als auch „echte“ Bipolar-Strukturen erzeugen kann, wird er als *BiCMOS*-Prozeß bezeichnet.

## 4.2 Der BiCMOS NPN-Transistor

Die Abbildung 4.8 zeigt das Schaltbild eines npn-Transistors mit dem Dioden-Ersatzschaltbild, mit Hilfe dessen man sich die pn-Übergänge des Transistors verdeutlichen kann.



**Abbildung 4.8:** Schaltungszeichen des npn-Bipolar Transistors und Dioden Ersatzschaltbild

Wie oben erwähnt, wird der hier zu behandelnde npn-Transistor mit einem anderen Prozeß hergestellt, weshalb andere Dotierprofile und andere Stoffe als Donatoren und Akzeptoren verwendet werden. Auch diese Dotierprofile wurden dem Asic-Labor Heidelberg, von der Firma *Austria Mikro Systeme International AG* zur Verfügung gestellt. Als npn-Transistor besteht dieses Bauteil aus einem n-dotierten Kollektor, p-dotierter Basis und n-dotierten Emitter.

Betrachtet man das Layout des BiCMOS NPN-Transistors (im folgenden als NPN-Transistor bezeichnet), das in dem linken Bild der Abbildung 4.9 dargestellt ist, so fallen zunächst die kleinen Skalen in der  $y$ -Richtung (also senkrecht zur Oberfläche) auf. Der pn-Übergang der Basis-Emitter Diode zum Beispiel befindet sich nur ca.  $60nm$  unterhalb der Oberfläche des Chips. Als neue Strukturen treten der sogenannte *N-burried layer* (wörtlich: „vergrabene N-Schicht“) und die beiden *Sinker* auf. Der rechte Teil der Abbildung zeigt einen Schnitt, der mitten durch den Emitter, von der Oberfläche bis in  $5\mu m$  Tiefe reicht. Aufgetragen ist



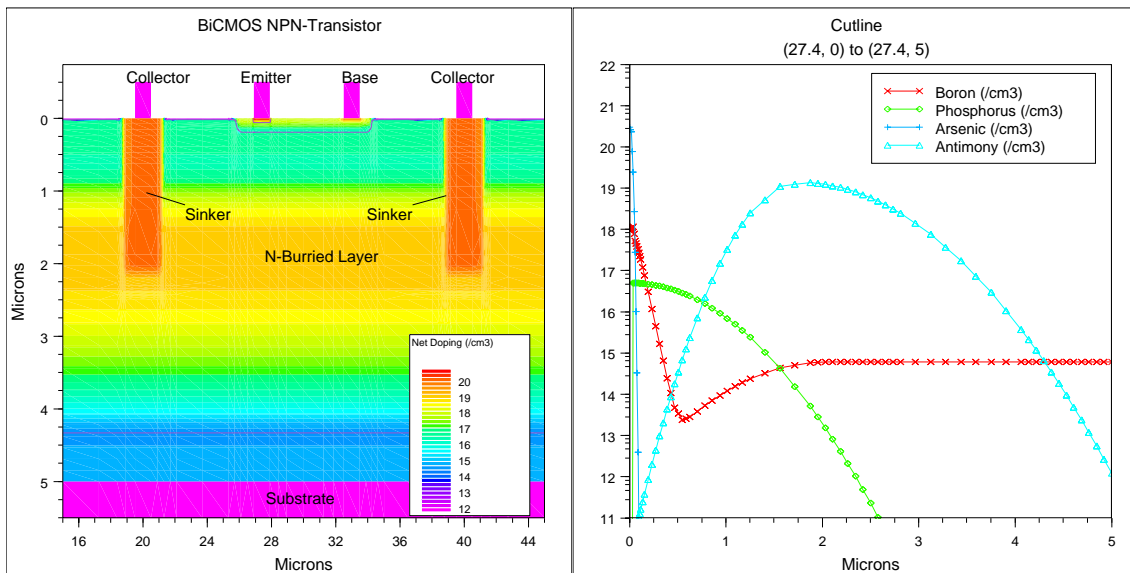


Abbildung 4.9: Layout und Dotierprofile des BiCMOS NPN-Transistors

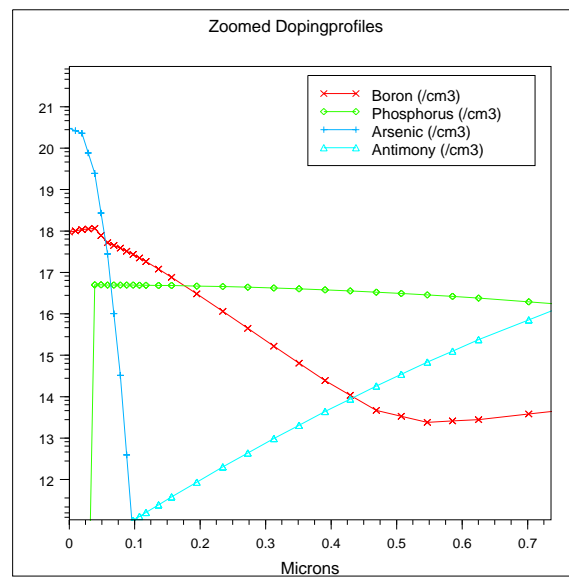


Abbildung 4.10: Dotierprofile des BiCMOS NPN-Transistors (Ausschnitt)

die Verteilung der Konzentration aller verwendeter Dotiermaterialien längs dieses Schnittes. Dabei sind zwei Stoffe hinzugekommen. Das sind Arsen (As) und Antimon (Sb), die, wie Phosphor beide in der fünften Hauptgruppe des Periodensystems der Elemente zu finden sind, und daher in Silizium als Donatoren wirken. Hier hat man also als Donatoren Phosphor, Arsen und Antimon und als Akzeptoren weiterhin nur Bor. Wie man in der Abbildung 4.10, die einen Ausschnitt des rechten Bildes der Abbildung 4.9 zeigt, ablesen kann, besteht der Emitter aus Arsen (dunkelblaue Kurve) mit einer maximalen Dichte von ca.  $2 \cdot 10^{20}/\text{cm}^3$ . Diese fällt sehr schnell ( $\approx 60\text{nm}$ ) um drei Größenordnungen ab. Danach folgt die Basis bestehend aus Bor, die, aus den im vorhergehenden Abschnitt erläuterten Gründen, relativ schwach dotiert ist (maximal  $1.2 \cdot 10^{18}/\text{cm}^3$ ). An der Stelle  $y \approx 160\text{nm}$  befindet sich der Basis-Kollektor Übergang, ab der dann die Phosphor-Dotierung des Kollektors die Netto-Dotierung bestimmt. Bei ca.  $0.7\mu\text{m}$  wird diese von der Antimon-Konzentration des n-burried layer dominiert, der wiederum hoch dotiert ist. Letztlich geht die Netto-Dotierung bei ca.  $4.3\mu\text{m}$  in die Bor-Dotierung des Substrates über.

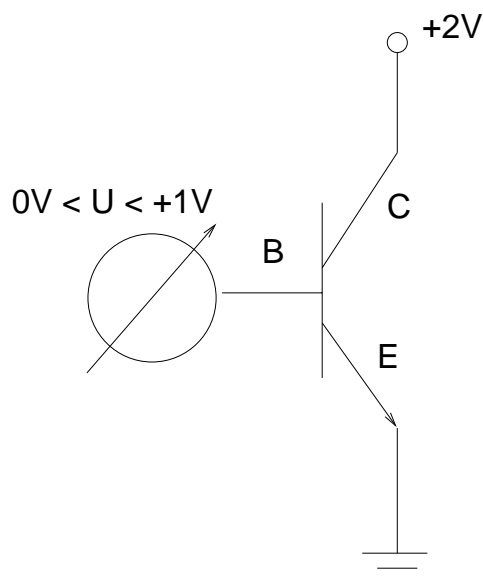
Wozu werden nun die Sinkers und der N-burried layer gebraucht?

Der hochdotierte n-burried layer dient dazu, unterhalb der Basis-Kollektor Übergangsschicht eine Äquipotentialfläche zu erzeugen, die parallel zur Oberfläche verläuft. Die Sinkers dienen nun dazu, das externe Potential in die Tiefe an den N-Burried Layer weiterzuleiten und diesen auf eben das an den Kollektorkontakten angelegte Potential zu bringen. Die beiden neuen Strukturen dienen also lediglich dazu die n-Dotierung des Kollektors, dotiert mit Phosphoratomen, zu kontaktieren. Das genaue Dotierprofil der Sinkers ist deshalb relativ unwichtig, da diese nur als niederohmige Verbindung gebraucht werden. Sie spielen ansonsten für die Funktion des Transistors keine Rolle. Diese Konfiguration ermöglicht es aber, verschiedene Transistoren mit eigenen Sinkern und N-burried layers, völlig getrennt zu betreiben. Das heißt insbesondere, daß die Kollektoren verschiedener Transistoren ihre eigenen externen Potentiale besitzen können und daß diese auch variiert werden können. Dies war beim Parasitären Vertikalen PNP-Transistor nicht möglich, da dort der Kollektor vom Substrat gebildet wurde, das auf dem kompletten Chip das selbe und feste Potential besitzt.

Um das Verhalten des NPN-Transistors des BiCMOS-Prozesses zu untersuchen, wird wieder die Common-Emitter Schaltung betrachtet.

Die Abbildung 4.11 zeigt die Schaltung mit den angelegten Spannungen. Sie sieht derjenigen vom vorhergehenden Abschnitt sehr ähnlich, die Potentiale haben aber, aufgrund der umgekehrten Schichtfolge an den pn-Übergängen, ihr Vorzeichen geändert.

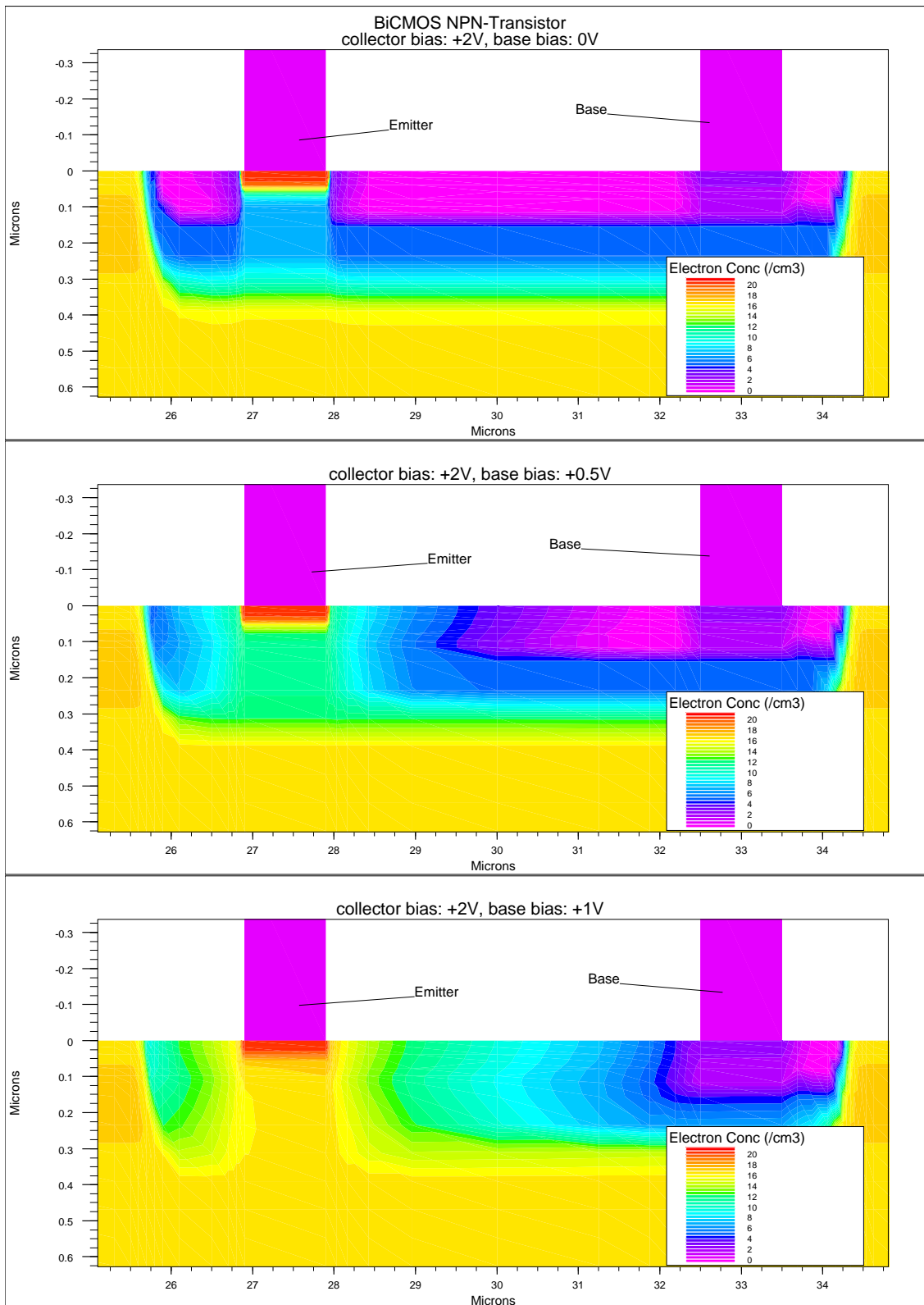
Die Basis ist beim npn-Transistor p-dotiert. Das heißt die Elektronen sind in der Basis die Minoritäten, weshalb in der Abbildung 4.12 die Elektronenkonzentration dargestellt ist. Zur besseren Sichtbarkeit zeigen die Bilder nur den Ausschnitt des Basisgebietes des Transistors. In allen drei Bildern beträgt die Kollektorspannung  $+2\text{V}$  und der Emitter ist an Masse gelegt ( $0\text{V}$ ). Im obersten Bild liegt an der Basis das externe Potential der Masse, zwischen Basis und Emitter fällt also keine Spannung ab, weshalb sich die Basis-Emitter Diode im thermodynamischen Gleichgewicht befindet. Dies kann man an der Raumladungszone am Emitter-Basis Übergang sehen. Im mittleren Bild ist die Basis-Emitter Diode mit  $0.5\text{V}$  Durchlaßspannung geschaltet, was sich deutlich am Anwachsen der Minoritätenkonzentration in der Basis bemerkbar macht. Schließlich ist im dritten Bild der Transistor durch eine Basis-Emitterspannung von  $+1\text{V}$  voll angesteuert und die Minoritätenkonzentration hat sich abermals drastisch erhöht. Die Abbildung 4.13 zeigt nochmals die Ladungsträgerkonzentrationen



**Abbildung 4.11:** Externe Spannungen der Common-Emitter Schaltung

in einem Schnitt mitten durch den Emitter bis in  $1\mu\text{m}$  Tiefe. Zunächst fällt im linken Bild auf, daß die Konzentration der Majoritäten in der Basis (Löcherkonzentration grün dargestellt) nicht mit der der Bor-Dotierung übereinstimmt. Das kommt daher, daß die Basis sehr schmal ist, und daß sich die Raumladungszone der Basis-Emitter Diode, die in diesem Bild ohne externe Potentialdifferenz geschaltet ist, aufgrund der sehr viel höheren Dotierung des Emitters vor allem in die Basiszone erstreckt. Weiterhin ist die Basis-Kollektordiode in Sperrichtung geschaltet, wodurch sich nochmals eine Raumladungszone ausbildet. Wie gesagt ist die Basis derart schmal, sodaß sich die beiden Raumladungszonen berühren, wodurch die Konzentration der Majoritäten in der Basis nicht der Dotierkonzentration entspricht. Da zwischen Basis und Emitter im linken Bild kein externes Potentialgefälle existiert, befindet sich dieser Übergang im thermodynamischen Gleichgewicht. Es gilt also das Massenwirkungsgesetz  $n \cdot p = n_i^2$ . Durch die oben beschriebene Absenkung der Majoritäten erhöht sich auch die Minoritätenkonzentration nach dem Massenwirkungsgesetz  $n = n_i^2 / p \approx 2.1 \cdot 10^{20} / (3 \cdot 10^{13} \text{ cm}^3) = 7 \cdot 10^6 / \text{cm}^3$ , wie man an der Skala ablesen kann. Im mittleren Bild ist durch die nun vorhandene Durchlaßspannung an der Basis-Emitter Diode die Raumladungszone geschrumpft, wodurch sich die Löcherkonzentration entsprechend der Dotierung in der Basis eingestellt hat. Die Elektronenkonzentration (rote Kurve) hat sich aber ebenfalls erhöht. Das Massenwirkungsgesetz kann nun nicht mehr angewendet werden, da sich der pn-Übergang nicht mehr im thermodynamischen Gleichgewicht befindet. Vielmehr werden nun Elektronen vom Emitter in die Basis injiziert, weswegen sich deren Konzentration erhöht. Im rechten Bild ist der Fall der Sättigung erreicht, und die Basis ist mit Elektronen überschwemmt, die dort eigentlich Minoritäten sind, nun aber eine Konzentration erreicht haben, die mit der der Majoritäten vergleichbar ist. Die Abbildung 4.14 zeigt den Basis- und Kollektorstrom in Abhängigkeit der Basissspannung bei festgehaltener Kollektorspannung von  $+2\text{V}$  und auf Masse gelegtem Emitter. Zusätzlich ist noch die Stromverstärkung  $\beta = I_C / I_B$  eingetragen, die in ihrem Maximum ca. 100 beträgt.

Der BiCMOS NPN-Transistor besitzt relativ kleine Strukturen. Der Emitter zum Beispiel



**Abbildung 4.12:** Elektronenkonzentration des BiCMOS NPN-Transistors bei verschiedenen externen Spannungen

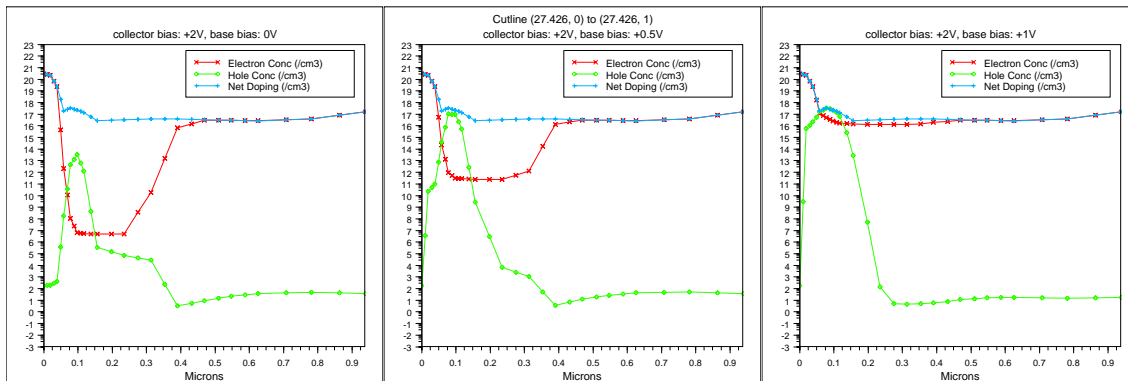


Abbildung 4.13: Ladungsträgerkonzentrationen des BiCMOS NPN-Transistors bei verschiedenen externen Spannungen

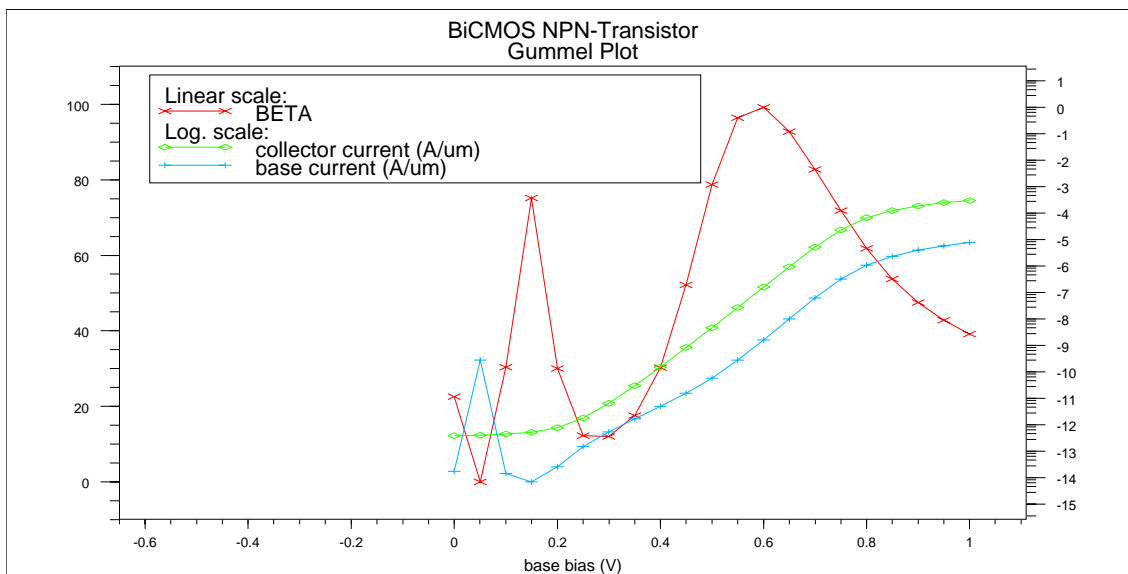
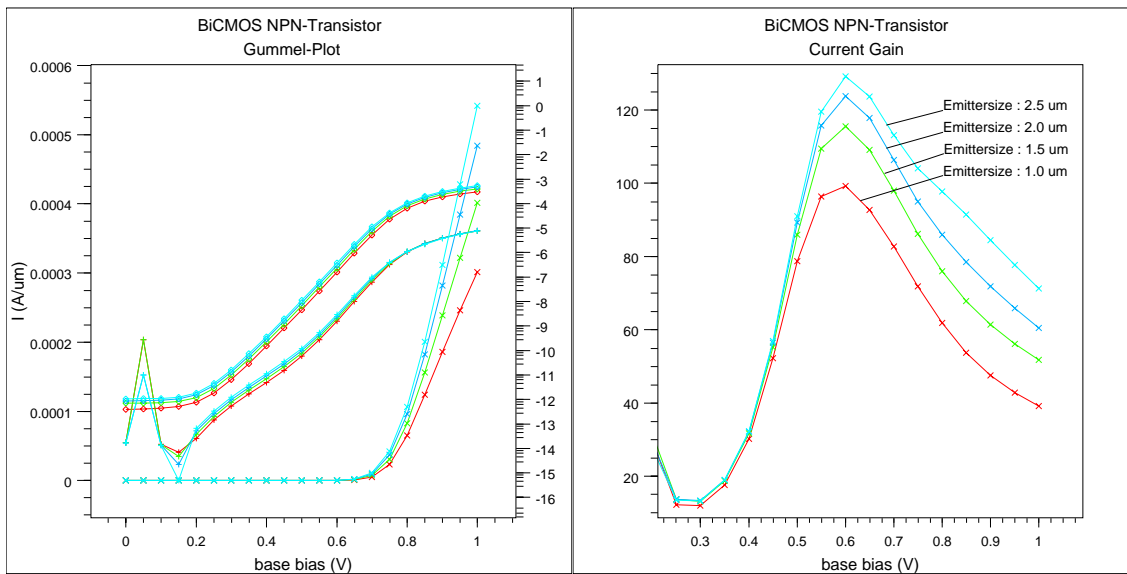


Abbildung 4.14: Basisstrom, Kollektorstrom und Stromverstärkung des BiCMOS NPN-Transistors

ist nur  $1\mu\text{m}$  breit. Es wurde darauf geachtet, die von der Chip-Herstellerfirma gelieferten Dotierprofile beim Design der Bauteile nachzubilden. Dies ist aber mit gewissen Fehlern verbunden. Die exakte Größe der Emittterbreite ist auch deshalb mit Unsicherheiten behaftet, da das seitliche Wegdiffundieren (im Simulator mit *Roll-Off* bezeichnet) der Dotiermaterialien während des Herstellungsprozesses ebenfalls die effektive Emitttergröße beeinflusst. Es war nicht immer möglich die Roll-Offs exakt den Herstellerdaten anzugleichen. Um zu verdeutlichen, wie Abweichungen davon die Eigenschaften des Bauteils beeinflussen, wurde der Gummel-Plot des BiCMOS NPN-Transistors mit vier verschiedenen Emitttergrößen simuliert. Das sind

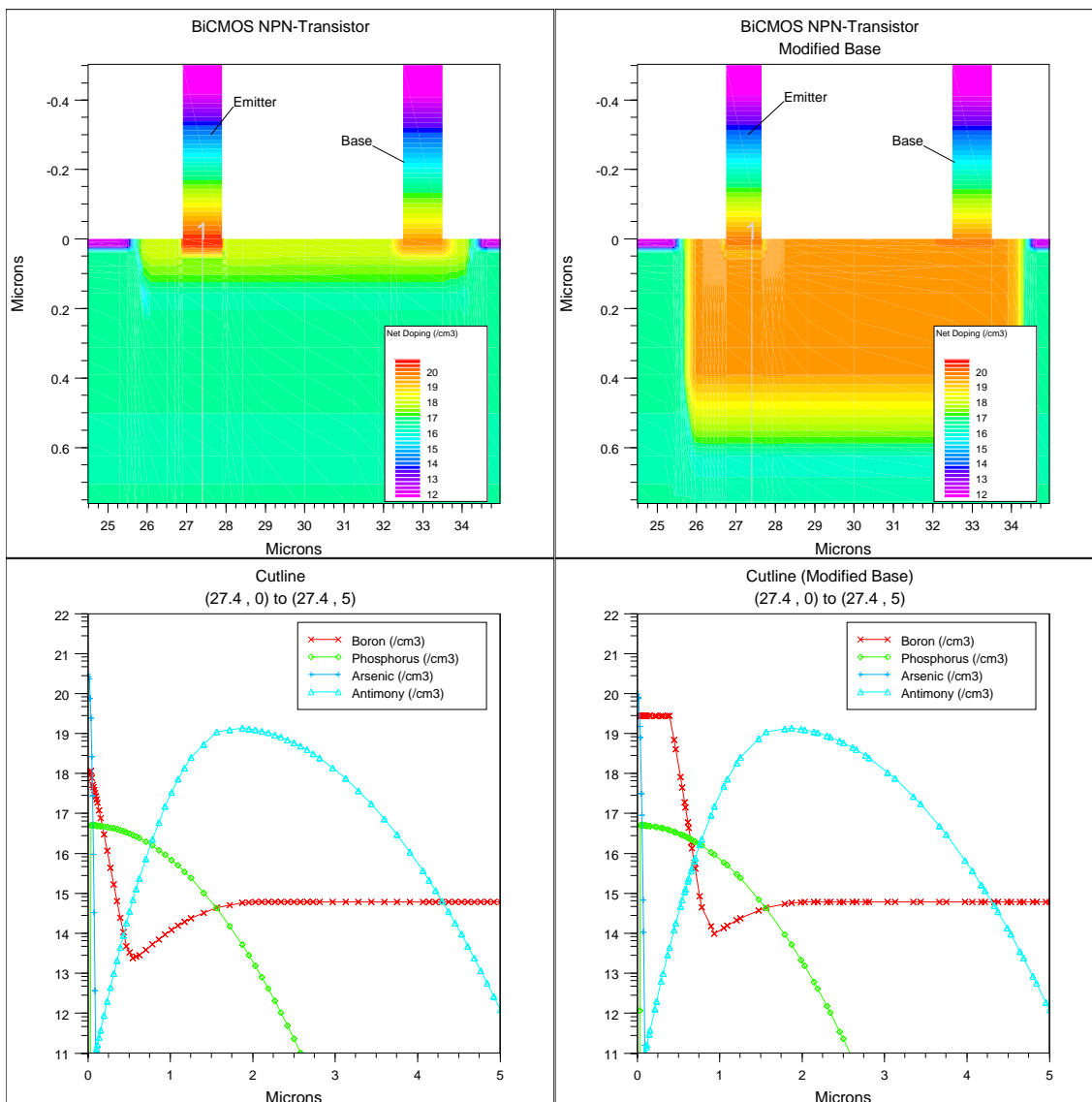


**Abbildung 4.15:** Gummelplot und Stromverstärkung des BiCMOS NPN-Transistors mit verschiedenen Emitttergrößen

die Originalgröße von  $1\mu\text{m}$ , und drei weitere, die jeweils einen um  $0,5\mu\text{m}$  breiteren Emittter besitzen. Die Abbildung 4.15 zeigt diesen Vergleich. In der linken Hälfte des Bildes sieht man in logarithmischer Skala den Basis- und Kollektorstrom. Wie man in der Abbildung 4.12 erkennen kann, ändern sich die Ladungsträgerkonzentrationen hauptsächlich unterhalb des Emittters, wenn der Transistor in verschiedenen Spannungszuständen betrieben wird. Es ist deshalb nicht verwunderlich, daß eine Änderung der Emitttergeometrie vor allem dessen elektrische Eigenschaften beeinflusst. Der Basisstrom bleibt nahezu unberührt, vor allem in dem Bereich, in dem die Basis-Emittter Diode ihre Schwellenspannung überschritten hat ( $U_B \geq 0,7\text{V}$ ). Um die Abweichungen im Kollektorstrom, der ja wesentlich vom Emittterstrom gebildet wird, besser zu sehen, wurde dieser zusätzlich in linearer Skala mit eingetragen. Die Erhöhung des Emittter- bzw. Kollektorstromes wirkt sich auf die Stromverstärkung des Transistors aus. Im rechten Bild ist die Stromverstärkung der Transistoren mit verschiedenen Emitttergrößen gezeigt. Die Farben entsprechen in beiden Bildern den gleichen Emitttergrößen.

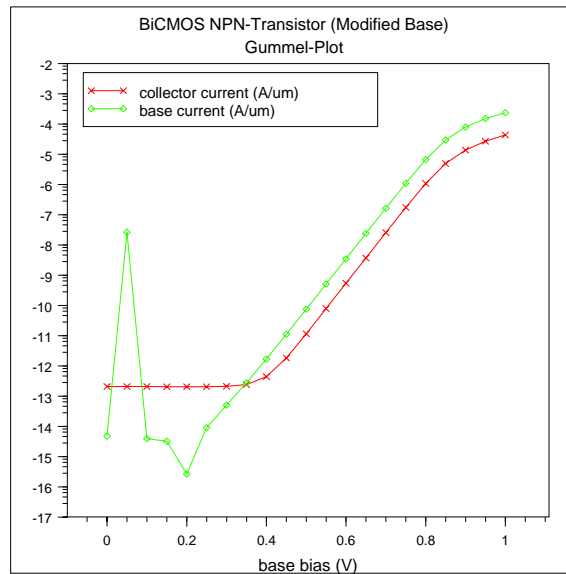
Ein weiterer Effekt soll noch kurz betrachtet werden. Im Abschnitt 4.1 wurde als Bedingung für das Zustandekommen der verstärkenden Eigenschaft des Transistors die relativ geringe Basisdotierung genannt. Es wird jetzt ein Bauteil gezeigt, das sich von dem des „normalen“ BiCMOS NPN-Transistors lediglich dadurch unterscheidet, daß die Basisdotierung modifiziert

worden ist. Die Abbildung 4.16 zeigt zum einen nochmals das Layout des ursprünglichen



**Abbildung 4.16:** Vergleich der Layouts zweier BiCMOS NPN-Transistoren mit verschiedenen Basisdotierungen

NPN-Transistors im Bereich der Basis und rechts daneben diejenigen mit der modifizierten Dotierung der Basis. Dabei ist folgendes zu beachten: erstens ist das Maximum der Dotierung um mehr als eine Größenordnung gestiegen, und zweitens hat sich die Basis verbreitert, so daß die Basis-Kollektor Übergangszone nun tiefer im Substrat liegt. Das sind also beides Ursachen für eine erhöhte Rekombination der, vom Emitter in die Basis injizierten, Elektronen mit den Majoritäten (Löcher) in der Basis. Das Resultat ist in der Abbildung 4.17 zu sehen. Dort sind der Basis- und Kollektorstrom in Abhängigkeit der angelegten Basissspannung aufgetragen. Der Kollektor ist, wie bei den anderen Gummel-Plots, auf ein externes Potential von  $+2V$  gelegt, während der Emitter auf Masse liegt. Die grüne Kurve stellt den Basisstrom dar. Er



**Abbildung 4.17:** Gummel-Plot des BiCMOS NPN-Transistors mit erhöhter Basisdotierung

überwiegt den Kollektorstrom um etwa eine Größenordnung. Die verstärkende Eigenschaft des Transistors ist durch die hohe Basisdotierung und die Vergrößerung der Basiszone zerstört worden. Daher erübrigt sich das Auftragen der Stromverstärkung  $\beta = I_C/I_B$ , da diese hier deutlich kleiner als eins ist.

Gegenüber dem Parasitären Vertikalen Bipolar Transistor hat der BiCMOS-Transistor den Vorteil, daß seine Gebiete vollständig vom Substrat isoliert sind, was bedeutet, daß die Kontakte, im Rahmen dessen was dem Transistor zuzumuten ist, mit beliebigen Potentialen versehen werden können. Er unterliegt also nicht den Einschränkungen, des Parasitären Vertikalen Bipolar Transistors, dessen Kollektor mit dem Substrat identisch ist und deshalb mit dem niedrigsten Potential der Schaltung verbunden werden muß und nicht variiert werden kann. Ein weiterer Vorteil des BiCMOS-Transistors ist seine um etwa das fünffache höhere Stromverstärkung verglichen mit der des Parasitären Vertikalen Bipolar Transistors. Der Nachteil des BiCMOS-Transistors allerdings sind seine Strukturen und Dotiermaterialien, die einen aufwendigeren Herstellungsprozeß benötigen, sodaß dieser Transistor nicht mit dem herkömmlichen und vor allem billigeren standardisierten CMOS-Prozeß hergestellt werden kann.



# Kapitel 5

## MOS-Transistor

Im vorhergehenden Kapitel wurden Bipolar-Transistoren untersucht. Sie werden so genannt, weil zum Stromfluß sowohl Elektronen als auch Löcher beitragen. In diesem Kapitel wird ein Vertreter der MOS-Transistoren vorgestellt. Diese werden auch als *unipolare* Transistoren bezeichnet, weil bei ihnen *entweder* Elektronen (NMOS) *oder* Löcher (PMOS) für den Fluß des elektrischen Stromes verantwortlich sind. Der Name der MOS-Transistoren leitet sich aus der Schichtfolge des Steueranschlusses ab. MOS steht für Metal Oxide Semiconductor. Er bezeichnet also die Schichtfolge Metall-Oxid-Halbleiter. Im folgenden wird ein NMOS-Transistor besprochen.

### 5.1 Der NMOS-Transistor

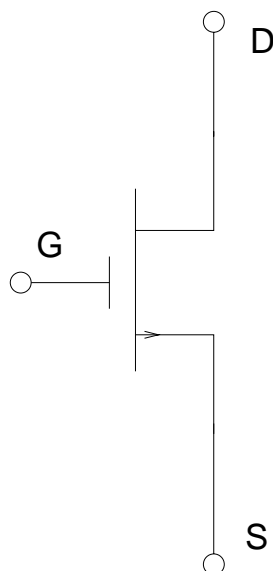


Abbildung 5.1: Schaltungszeichen des NMOS-Transistors

Die Abbildung 5.1 zeigt das Schaltbild eines NMOS-Transistors. Wie beim Bipolar Transistor gibt es auch hier drei elektrische Kontakte, die als *Source*, *Gate* und *Drain* bezeichnet werden. Die Funktionen entsprechen denen des Bipolar-Transistors. Das Gate (wörtlich: Tor) ist der Steueranschluß und entspricht der Basis, die Source (wörtlich: Quelle) dem Emitter und schließlich der Drain (wörtlich: Abfluß) dem Kollektor. Im Schaltzeichen ist am Gate-Anschluß angedeutet, daß der Gate-Kontakt nicht direkt mit dem Halbleiter in Verbindung steht, sondern durch eine Isolierschicht von diesem getrennt ist (daher auch der Name MOS). Über das Gate wird deshalb im statischen Betrieb (bis auf Leckströme) kein Strom fließen. Der Strom, der von Drain nach Source fließt, wird durch das elektrische Feld, welches sich durch die Isolierschicht vom Gate auf die Halbleiteroberfläche fortpflanzt, gesteuert. Die Transistoren werden daher auch MOSFET's (Field Effect Transistor) genannt. Im linken Bild der

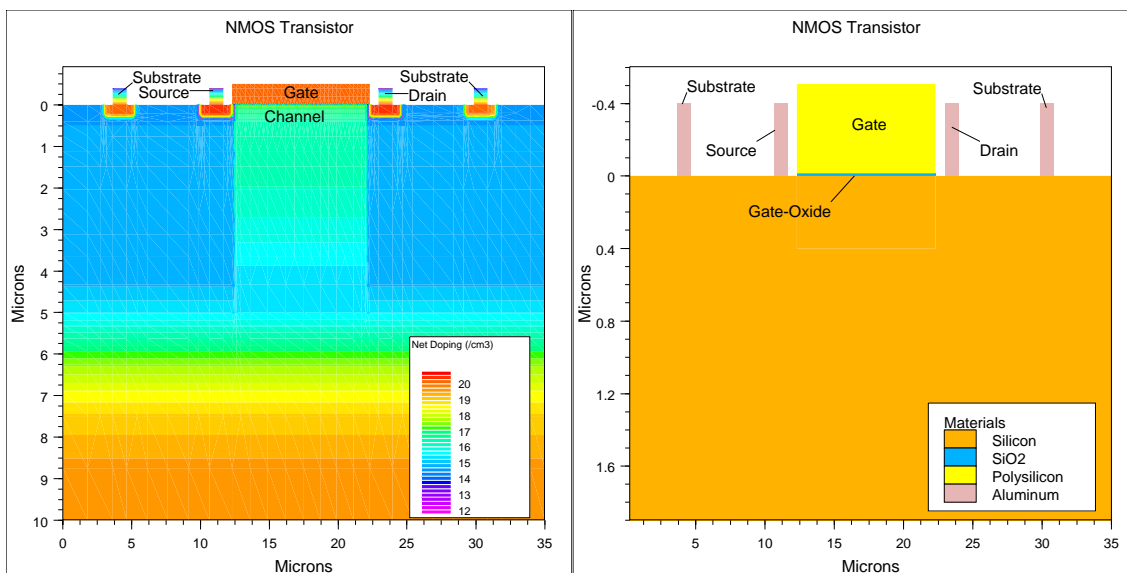
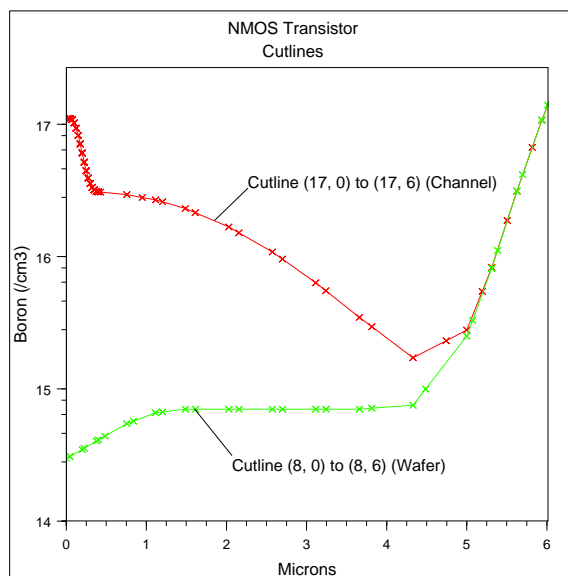


Abbildung 5.2: Layout und Schichtfolge des NMOS-Transistors

Abbildung 5.2 ist die Netto-Dotierung dargestellt. Das Substrat ist p-dotiert. Um diesem ein definiertes Potential zu geben (üblicherweise  $0V$ , in jedem Falle aber das niedrigste Potential der Schaltung) sind zur Ankontaktierung unterhalb der Substrat Aluminium-Kontakte rechts und links des Transistors hochdotierte p-Zonen angebracht. Die Gebiete unterhalb von Drain und Source sind n-Diffusionen, die das gleiche Tiefenprofil haben, wie die n-Diffusion der N-Diffusion/Substrat Diode aus Kapitel 3 (diese Dioden wurden ja gerade mit den Dotierprofilen des CMOS-Prozesses erzeugt). Als neues Dotierprofil kommt das unterhalb der Kanalregion (in der Abbildung 5.2 mit „Channel“ markiert) hinzu. Die Region wird zusätzlich mit Akzeptoren (hier: Bor) dotiert. In der Abbildung 5.3 sind zwei Dotierprofile zu sehen, die sich bis ca.  $5\mu\text{m}$  Tiefe unterscheiden und dann identisch sind. Die grüne Kurve ist bekannt und beschreibt die Dotierung des Wafers. Sie dient zum Vergleich mit der neu hinzugekommenen Dotierung der Kanalregion (rot). Wie später noch zu sehen ist, hat die Kanaldotierung maßgeblichen Einfluß auf die Funktion des NMOS-Transistors, insbesondere auf die sogenannte *Threshold-Spannung*. Im rechten Bild der Abbildung 5.2 sind durch farbliche Markierung die benutzten Materialeien zum Aufbau des NMOS-Transistors gezeigt. Der Wafer, der aus Silizium besteht, ist orange markiert. Die Metallkontakte bestehen aus Aluminium. Oben

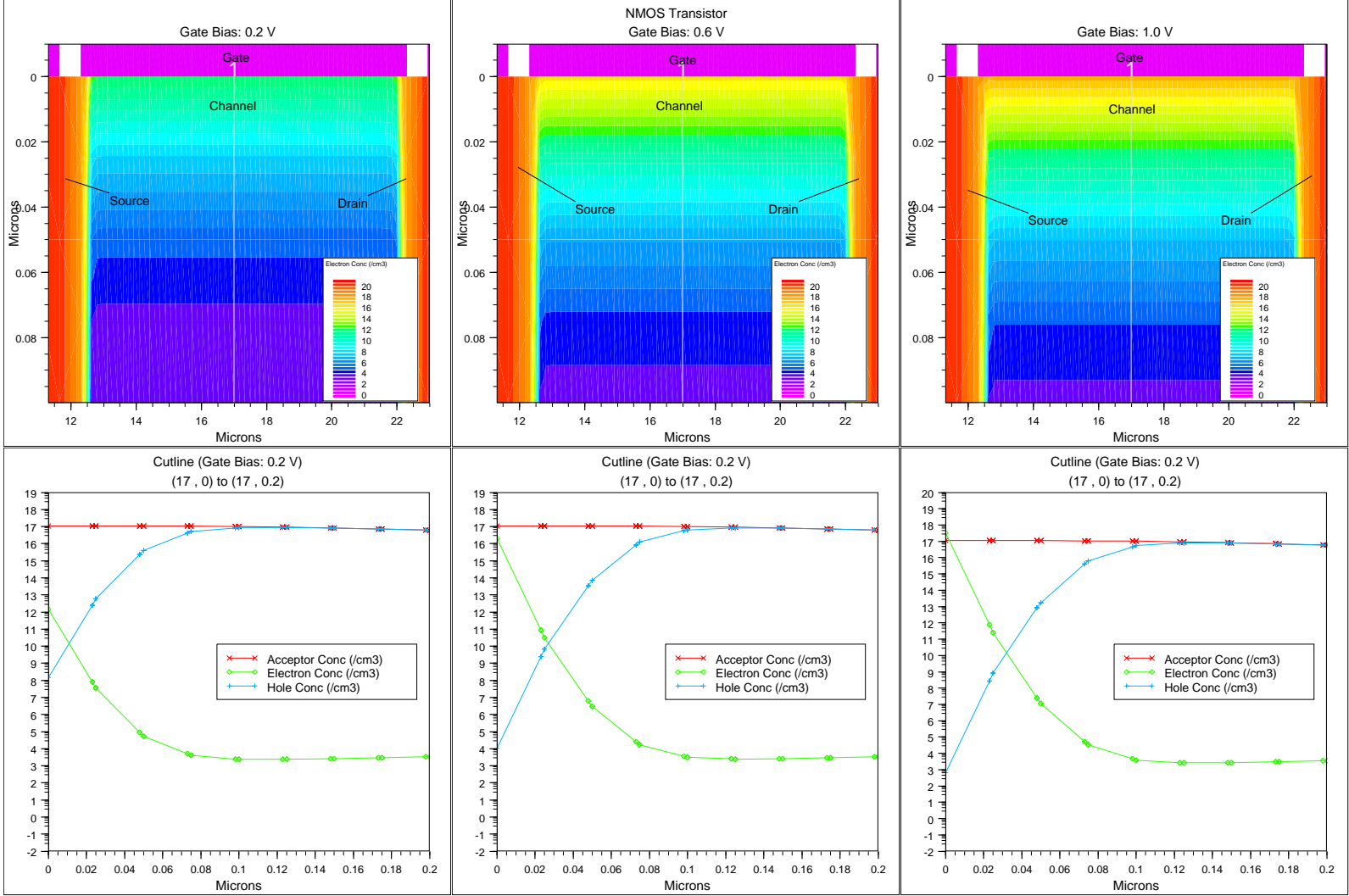


**Abbildung 5.3:** Schnitt durch die Kanalregion und durch das Substrat

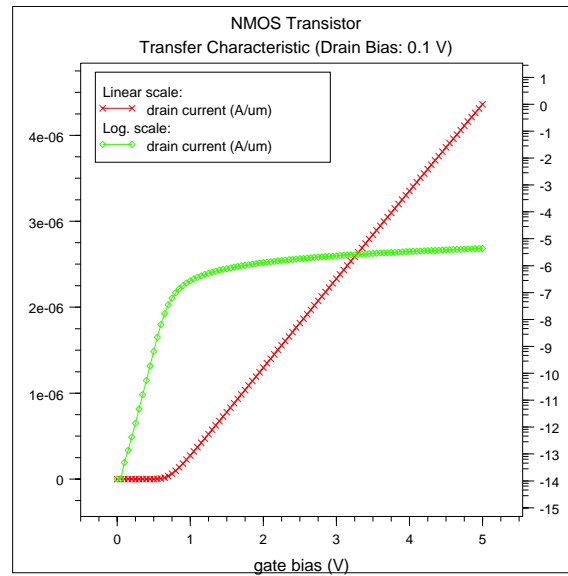
wurde der Name des NMOS durch seine Schichtfolge am Gate erklärt. Hier sieht man aber, daß der Gatekontakt nicht aus Metall besteht (NMOS), sondern aus Polysilizium, das, wie man am linken Teil der Abbildung 5.2 erkennen kann, hoch dotiert ist. Durch die hohe n-Dotierung hat es metallische Eigenschaften. Zwischen dem Gate-Kontakt aus Polysilizium und dem Halbleitermaterial befindet sich eine sehr dünne Schicht aus Siliziumdioxid ( $SiO_2$ ), welche die Rolle der Isolierschicht übernimmt (NMOS). Diese Schicht ist nur  $16\text{nm}$  dick und isoliert das Gate vom Kanal.

Um die Wirkungsweise des NMOS zu verdeutlichen, ist in der Abbildung 5.4 ein Ausschnitt des NMOS gezeigt. Es ist nur die Kanalregion zu sehen, an der sich rechts und links das Source- bzw. Drain-Gebiet anschließt. Man beachte hier die extrem in y-Richtung gedehnte Skala der Längenangaben. Die unteren drei Bilder zeigen einen Schnitt durch diese Region, wobei außer der Elektronenkonzentration (grün) und der Löcherkonzentration (blau) zusätzlich die Bor-Dotierung (rot) mit eingezeichnet ist. Man sieht, daß die Löcherkonzentration, die unter normalen Umständen der Bor-Dotierung entsprechen müßte (mit der Näherung  $N_A \approx N_A^-$ ), in der Kanalregion deutlich erniedrigt ist. Da dem Transistor von außen keine Ladung zugeführt wird (der Gate-Kontakt ist isoliert, die anderen Spannungen betragen  $0\text{V}$ ), gilt das Massenwirkungsgesetz  $n \cdot p = n_i^2 = \text{const.}$  Das bedeutet für die Elektronen, daß ihre Dichte dort erhöht wird. Die „normalen“ Ladungsträgerverhältnisse werden erst in ca.  $120\text{nm}$  Tiefe erreicht, wo das Gate-Potential keinen Einfluß mehr auf die Ladungen hat. Näher an der Oberfläche ist die Elektronendichte aufgrund des Potentials am Gate deutlich angestiegen. Es hat sich ein n-leitender *Kanal* ausgebildet, der mit zunehmender Gate-Spannung größer wird. Anschaulich kann man sich vorstellen, daß aufgrund eines positiven Gate-Potentials die Löcher aus der Kanalregion „rausgedrückt“ bzw. die Elektronen von diesem „angezogen“ werden. Wird nun zwischen Drain und Source eine Spannung angelegt, so können diese angereicherten Ladungsträger über den Kanal von Drain nach Source fließen. Der Strom zwischen Source und Drain wird nun durch die Gate-Spannung gesteuert, indem diese je nach

Abbildung 5.4: Elektronenkonzentration des NMOS bei verschiedenen Gate-Spannungen



Höhe der Spannung eine höhere oder niedrigere Elektronenkonzentration in der Kanalregion hervorruft. Um zu sehen, wie das Gatepotential den Strom von Drain nach Source steuert, ist in der Abbildung 5.5 der Drain-Strom in Abhängigkeit der Gate-Spannung aufgetragen. Diese Art der Kennlinie nennt man Übertragungskennlinie, denn sie vermittelt einen Zusammenhang zwischen der Eingangsgröße (Gate-Spannung) und Ausgangsgröße (Drain-Strom). Sie entspricht dem Gummel-Plot der Bipolar Transistoren, mit dem Unterschied, daß hier, aufgrund der Isolation des Gates, durch die Oxidschicht kein Gate-Strom fließen kann. Die Drain-Spannung ist hier auf  $0.1V$  gelegt, Source und Substrate liegen auf Masse. Dabei fällt



**Abbildung 5.5:** Übertragungskennlinie des NMOS mit  $V_D = 0.1V$

zunächst auf, daß der NMOS ebenfalls so etwas wie eine Schwellenspannung hat. Betrachtet man den Drainstrom in linearer Skala (rot), so sieht man, daß der Drainstrom erst ab einer Gate-Spannung von ca.  $0.65V$  anfängt, stark anzuwachsen. Man nennt diese Spannung *Threshold-Spannung* (wörtlich: Schwellenspannung). An der logarithmischen Skala kann man sehen, daß der Strom unterhalb der Threshold-Spannung (sogenannter *Sub-Threshold Bereich*) logarithmisch ( $\log(I_D) \propto V_G$ ) ansteigt, dann aber abknickt und nun der Drainstrom nur noch proportional zu  $V_G^\kappa$  steigt, wobei  $\kappa \approx 4/3$  beträgt. Es gibt verschiedene Definitionen für die Threshold-Spannung: nähert man auf der linearen Skala den Drainstrom im Bereich über der Thresholdspannung einer Geraden an, so markiert der Durchstoßpunkt der Geraden mit der Konstanten  $y = 0$  die Threshold Spannung. Mit diesem Verfahren ergibt sie sich hier zu  $\approx 0.72V$ . Bei dieser Näherung sollte die Drain-Source Spannung kleiner  $0.1V$  sein.

Im folgenden sollen die Ausgangskennlinien des NMOS betrachtet werden. Dazu dient die Gate-Spannung als Parameter, die jeweils auf einen festen Wert eingestellt wird, das Source-Potential bleibt auf  $0V$  und die Drainspannung wird hochgefahren. Die Abbildung 5.6 zeigt das Kennlinienfeld bei Gate-Spannungen von  $1V$ ,  $2V$ ,  $3V$  und  $4V$ . Die Skala ist linear und man erkennt einen ohm'schen Bereich ( $I_D \propto U_D$ ) bei kleinen Drain-Spannungen. Die Kennlinien knicken ab einem bestimmten Wert der Drain-Spannung ab, und der Transistor befindet sich nun im Sättigungsbereich. Es liegt also kein ohm'sches Verhalten mehr vor.

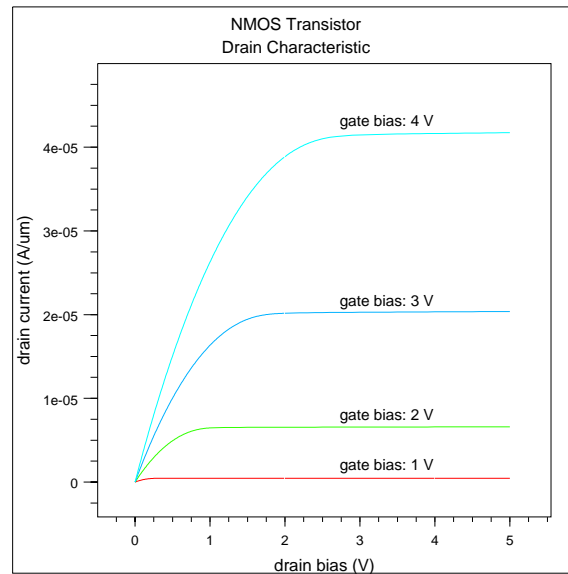


Abbildung 5.6: Ausgangskennlinienfeld des NMOS

Ein Grund dafür ist die abnehmende Beweglichkeit  $\mu_{n,p}$  der Ladungsträger mit ihrer zunehmenden Konzentration. Wie in Kapitel 2 beschrieben nehmen Streuprozesse der Ladungsträger untereinander zu, wenn sich die Ladungsträgerdichten erhöhen. Die Geschwindigkeit der Ladungsträger ist dann nicht mehr proportional zum beschleunigenden Feld, wie das im ohm'schen Bereich der Fall ist ( $\vec{v} = \mu_{n,p} \vec{E}$ ). Dies gilt nur solange  $\mu_{n,p} = \text{const.}$  Mit zunehmenden Feld (hier ist das longitudinale Feld  $\vec{E}_{||}$ , hervorgerufen durch das Drain-Potential, gemeint) sinkt die Beweglichkeit der Ladungsträger und es gilt  $\mu_{n,p} = \mu_{n,p}(\vec{E}_{||})$ . Die Geschwindigkeit der Ladungsträger geht gegen eine Sättigungsgeschwindigkeit, ab der sich ein weiteres Anwachsen des beschleunigendes Feldes nicht mehr in einer Zunahme der Geschwindigkeit bemerkbar macht.

Ein weiterer Grund ist die sogenannte *Kanallängenmodulation*. Darunter versteht man die abnehmende Kanallänge mit zunehmender Drain-Spannung. Sie wird auch als *Kanalabschnürung* (engl.: pinch-off) bezeichnet. Das Drain- bzw. Source-Gebiet besteht aus jeweils einer n-Diffusion im p-Substrat. Diese Konfiguration wurde im Kapitel 3 als N-Diffusion/Substrat Diode vorgestellt. Dort wurde auch gezeigt, daß eine Sperrspannung am pn-Übergang sich in einer Ausweitung der Raumladungszone bemerkbar macht. Betrachtet man das Ausgangskennlinienfeld, so erkennt man, daß der pn-Übergang, bestehend aus n-Diffusion des Drain-Gebietes und p-Dotierung des Substrates, mit zunehmender Drain-Spannung mehr und mehr in Sperrichtung geschaltet wird. Das bedeutet aber, daß sich um das Drain-Gebiet eine immer größer werdende Raumladungszone, die ja frei von beweglichen Ladungsträgern ist, ausbildet. Diese erstreckt sich auch in die Kanalregion hinein. Die vom Gate-Potential in die Kanalregion „angesaugten“ Elektronen werden daher im Bereich der Raumladungszone, der in Sperrichtung gepolten Drain/Substrat Diode, mit zunehmender Sperrspannung (also positiver werdendem Drain-Potential) dem wachsenden Feld der Raumladungszone unterworfen. Dieses Feld ist so gerichtet, daß es die Elektronen der p-Seite, die dort normalerweise Minoritäten sind, in das n-Gebiet zieht. Das geschieht dann aber auch mit den Elektronen

in der Kanalregion, die dem Kanal dann nicht mehr zur Verfügung stehen und den Kanal hochohmiger werden lassen

Ein weiterer Grund ist die zunehmende Abschirmung des transversalen Feldes  $\vec{E}_\perp$ . Die Raumladungszone der in Sperrichtung gepolten Drain/Substrat Diode hinterläßt auf der p-Seite des Übergangs negativ ionisierte Akzeptoren ( $N_A^-$ ), deren Zahl und räumliche Ausdehnung mit wachsender Sperrspannung zunehmen. Diese negativen Ladungen schirmen das positive Potential der Gate-Spannung vom Substrat unterhalb der Kanalregion ab. Dadurch wird verhindert, daß sich Influenzladungen, bestehend aus freien Elektronen, in der Kanalregion ansammeln, wodurch sich deren Konzentration verringert. Dies wiederum erhöht den elektrischen Widerstand des Kanals und hindert so die Zunahme des Drainstromes. In der Abbildung

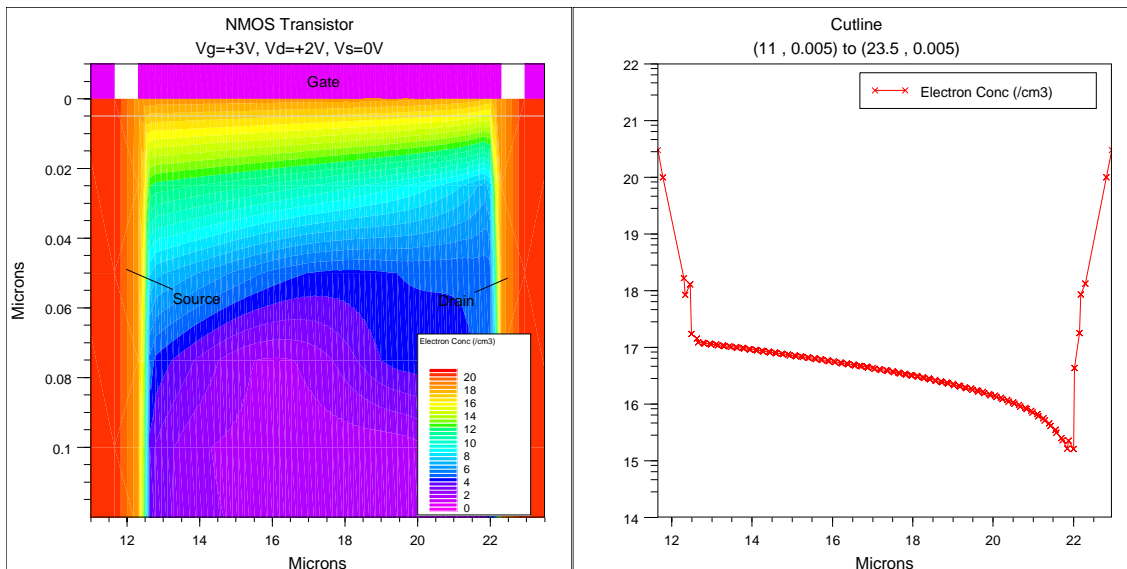
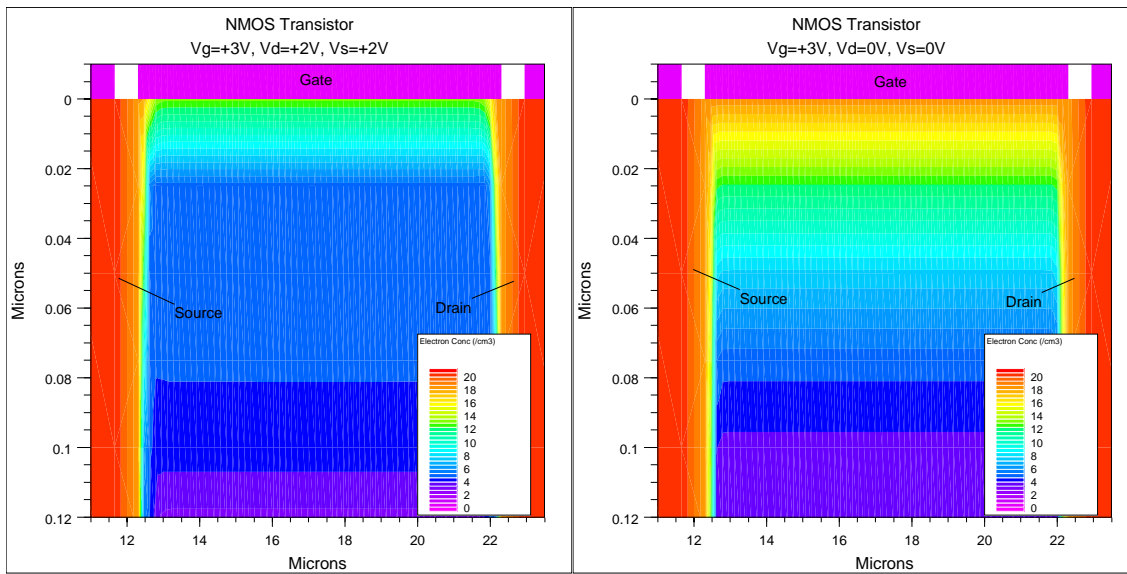


Abbildung 5.7: Kanalabschnürung bei  $V_D = +2V$

5.7 kann man den Effekt der Kanalweitenmodulation erkennen. Die Gate-Spannung beträgt  $+3V$ , die Drain-Spannungen  $+2V$  und der Source-Kontakt ist auf Masse gelegt. Das sind die Potentiale, die in Abbildung 5.6 der dunkelblauen Kurve am Anfang des Sättigungsgebietes (drain bias:  $2V$ ) entsprechen. An der Elektronenkonzentration in der Kanalregion kann man deutlich sehen, daß sich nahe des Drain-Gebietes die Dichte stark verringert hat. Im rechten Teil der Abbildung ist ein Schnitt parallel zur Oberfläche in  $5nm$  Tiefe längs des Kanals gemacht und die Elektronenkonzentration aufgetragen. Links und rechts des Bildes steigt sie gemäß der hohen Dotierung von Source und Drain an. Die Elektronenkonzentration beträgt auf der Drain-Seite nur noch etwa 1.2% des Maximums auf der Source-Seite. In manchen Büchern ist zu lesen, daß die Verengung des Kanals durch den Spannungsabfall längs des Kanals aufgrund des Drain-Source Potentialgefälles verursacht wird. Im linken Bild der Abbildung 5.8 wurde deshalb, bei einer Gate-Spannung von  $+3V$ , Source und Drain auf das gleiche Potential, nämlich  $+2V$  gelegt, so daß zwischen Drain und Source kein Spannungsabfall existiert, und somit auch kein Strom fließen kann. Wie man aber an der Elektronenkonzentration ablesen kann, wird der Kanal trotzdem erheblich verengt. Dies geschieht aufgrund der Spiegelsymmetrie dieses NMOS-Transistors ebenfalls symmetrisch. Das rechte Bild dient zum Vergleich: die Gate-Spannung beträgt ebenfalls  $+3V$ , die Source- und Drain-Spannung



**Abbildung 5.8:** Vergleich der Elektronenkonzentration in der Kanalregion bei verschiedenen Source und Drain Potentialen.

aber  $0V$ .



## Kapitel 6

# Optische Simulationen

Der Simulator bietet die Möglichkeit, elektromagnetische Wellen im visuellen Bereich auf die Bauteile einzustrahlen. Dazu muß die Lichtquelle definiert werden. Sie wird festgelegt durch:

- die Wellenlänge des einfallenden Strahles,
- den Ursprung der Quelle in x- und y-Koordinaten,
- den Einfallswinkel des Strahles auf die Oberfläche des Bauelementes,
- die Breite des einfallenden Strahles,
- die Intensität des einfallenden Strahles,
- die Anzahl der Reflexionen des Strahles an Grenzschichten.

An der Oberfläche des Bauelementes splittet sich nach der klassischen Optik der einfallende Strahl in einen reflektierten und transmittierten Strahl auf. Die Intensitäten des reflektierten ( $E_r$ ) und transmittierten ( $E_t$ ) Strahles werden für parallel polarisiertes Licht nach folgenden Formeln berechnet:

$$E_r = \frac{n_1 \cos(\Theta_t) - n_2 \cos(\Theta_i)}{n_1 \cos(\Theta_t) + n_2 \cos(\Theta_i)} \cdot E_i$$

$$E_t = \frac{2n_1 \cos(\Theta_i)}{n_1 \cos(\Theta_t) + n_2 \cos(\Theta_i)} \cdot E_i.$$

Für senkrecht polarisiertes Licht gelten folgende Beziehungen:

$$E_r = \frac{n_1 \cos(\Theta_i) - n_2 \cos(\Theta_t)}{n_1 \cos(\Theta_i) + n_2 \cos(\Theta_t)} \cdot E_i$$

$$E_t = \frac{2n_1 \cos(\Theta_i)}{n_1 \cos(\Theta_i) + n_2 \cos(\Theta_t)} \cdot E_i.$$

Die Reflexions- und Transmissionskoeffizienten ergeben sich zu:

$$R = \left( \frac{E_r}{E_i} \right)^2$$

$$T = \left( \frac{E_t}{E_i} \right)^2 \frac{n_2}{n_1}$$

Die Winkel ergeben sich aus:

$$\Theta_r = \Theta_i$$

$$n_1 \sin \Theta_i = n_2 \sin \Theta_t.$$

Dabei bezeichnen:

$E_i$  die einfallende (engl: incident) Intensität,

$E_r$  die reflektierte Intensität,

$E_t$  die transmittierte Intensität,

$\Theta_i$  den Einfallswinkel gemessen zur Normalen der Halbleiteroberfläche,

$\Theta_r$  den Reflexionswinkel,

$\Theta_t$  den Transmissionswinkel,

$n_1, n_2$  den Brechungsindex des Medium 1 bzw. Medium 2  $R, T$  den Reflexions- bzw Transmissionskoeffizienten.

Die Photogenerationsrate  $G$  im Halbleiter wird durch

$$G = \eta_0 \frac{P^* \lambda}{hc} \alpha e^{-\alpha y}$$

gegeben. Dabei bezeichnen

$P^*$  die Modifikation der Intensität durch Reflexion, Transmission und Absorption,

$\eta_0$  die interne Quanteneffizienz, die die Zahl der Elektronen-Loch Paare pro absorbiertem Photon angibt,

$y$  die hinterlegte Strecke des Strahles im Medium,

$h, c$  die Planck'sche Konstante und die Lichtgeschwindigkeit,

$\lambda$  die Wellenlänge des Lichtes.

Der Absorptionskoeffizient  $\alpha$  wird durch

$$\alpha = \frac{4\pi}{\lambda} \cdot k$$

gegeben, wobei  $k$  der Imaginärteil des Brechungsindex des verwendeten Materials ist. Der Real- und Imaginärteil des Brechungsindex sind abhängig von der Wellenlänge des einfallenden Lichtes und werden vom Simulator-Hersteller aus Tabellen des „Handbook of Optical Constants“ interpoliert (siehe auch [8]).

Mehrfachreflexionen an Schichten (zum Beispiel einer Passivierungsschicht auf der Halbleiteroberfläche) werden vom Simulator berücksichtigt. Das drückt sich in der oben genannten Größe  $P^*$  aus. Es werden aber davon nur die Intensitäten der Lichtstrahlen berührt und keine Interferenzen berechnet. Effekte wie destruktive oder konstruktive Interferenz an dünnen Schichten (Fabry-Perot Filter o.ä.) werden nicht simuliert. Das Licht wird im Rahmen der klassischen Optik behandelt und dessen Wellennatur nicht berücksichtigt. In den nächsten Abschnitten werden verschiedene optische Simulationen mit den vorher beschriebenen Bauelementen durchgeführt. Dabei werden, wenn nicht anders angegeben, die Intensitäten des Lichtes auf  $1W/m^2$  eingestellt. Das entspricht ungefähr der Intensität bei Raumlicht. Die Wellenlängen wurden mit Rücksicht auf den im ASIC-Labor Heidelberg zur Verfügung stehenden Laser zu  $675nm$  gewählt. Die Breite des einfallenden Strahles wird  $10\mu m$  betragen. Dieser Lichtstrahl wird im folgenden als Standard-Spot bezeichnet. Die Bezeichnung Spot ist etwas irreführend. Man muß bedenken, daß der Simulator ein 2D-Simulator ist. Das bedeutet, daß alle Angaben auf eine Dimension normiert ist (z.B.  $A/\mu m$ ). Das bedeutet aber auch für den Spot, daß dieser nicht rund ist, wie man dem Namen entnehmen könnte, sondern ein rechteckiges Profil besitzt. Wenn also im folgenden vom Lichtstrahl oder Spot die Rede ist, so wird ein rechteckiges Profil definiert.

## 6.1 Photostrom in Abhängigkeit der angelegten Spannungen

Zunächst werden die N-Diffusion/Substrat Diode, N-Wanne/Substrat Diode und die P-Diffusion/N-Wanne Diode mit dem oben spezifizierten Licht-Spot beleuchtet und der Photostrom in Abhängigkeit der anliegenden Sperrspannung gemessen.

### 6.1.1 N-Diffusion/Substrat Diode

Die Abbildung 6.1 zeigt im linken Teil den Kathodenstrom der Diode in Abhängigkeit der Kathodenspannung. Im rechten Teil der Abbildung ist die Photo-Generationsrate bei einer Sperrspannung von  $5V$  gezeigt. Die senkrechten roten Striche sollen den Spot andeuten. Er ist in Wirklichkeit  $10\mu m$  breit (erstreckt sich also von  $x = -5\mu m$  bis  $x = +5\mu m$ ), wurde hier aber nicht in voller Breite eingezeichnet, um den darunter liegenden Kontur-Plot der Photo-Generationsrate nicht zu überdecken. Aufgrund der Absorption nimmt die Generationsrate mit zunehmender Tiefe ab.

Die einfallenden Photonen erzeugen im Halbleiter Elektronen-Loch Paare. Um zu verdeutlichen, wie das Licht die Ladungsträgerkonzentration im Bauteil beeinflusst, sind in der Abbildung 6.2 die Dichten der Ladungsträger bei verschiedenen Beleuchtungen gegenübergestellt. In beiden Bildern beträgt die Kathodenspannung  $+2V$ , die Diode ist also mit einer Sperrspannung von  $2V$  geschaltet. Im linken Bild wird kein Licht eingestrahlt, im rechten der oben beschriebene Standard-Spot. Dabei ist zu erkennen, wie sich die Ladungsträgerdichten aufgrund der einfallenden Photonen geändert haben. Von links kommend ( $x \geq 0\mu m$ ) durchquert man das n-Gebiet der Kathode. In diesem sind die Elektronen Majoritäten (rote Kurve). Aufgrund ihrer hohen Dotierkonzentration kann das Licht ihre Dichte nicht merkbar erhöhen. Betrachtet man dagegen in der selben Region die Löcher, die hier die Minoritäten sind, so erkennt man eine deutliche Erhöhung der Konzentration gegenüber dem linken Bild,

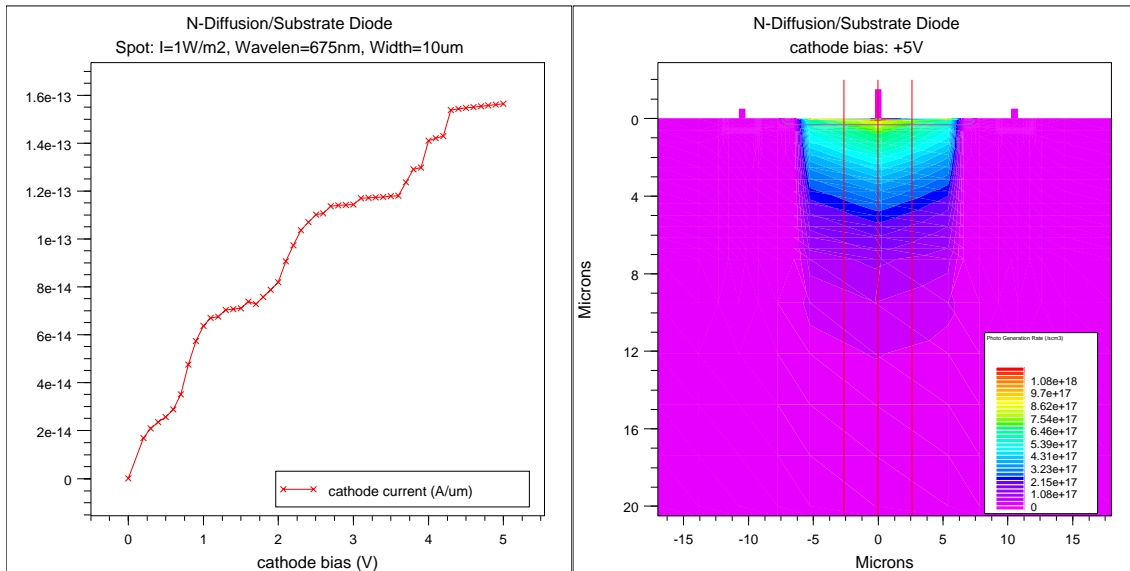


Abbildung 6.1: Kathodenstrom in Abhängigkeit der Kathodenspannung

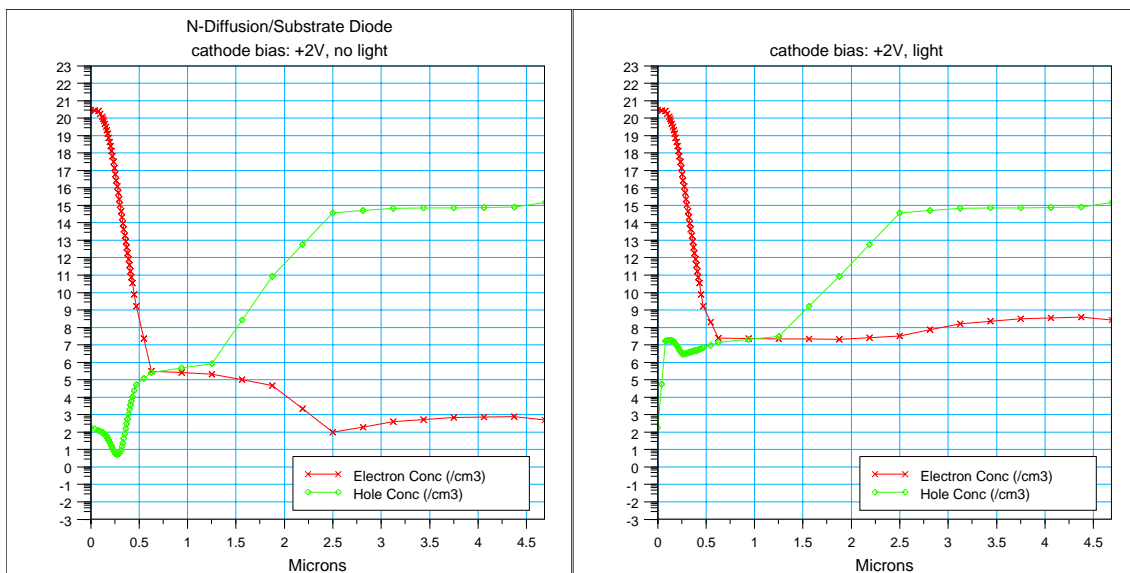
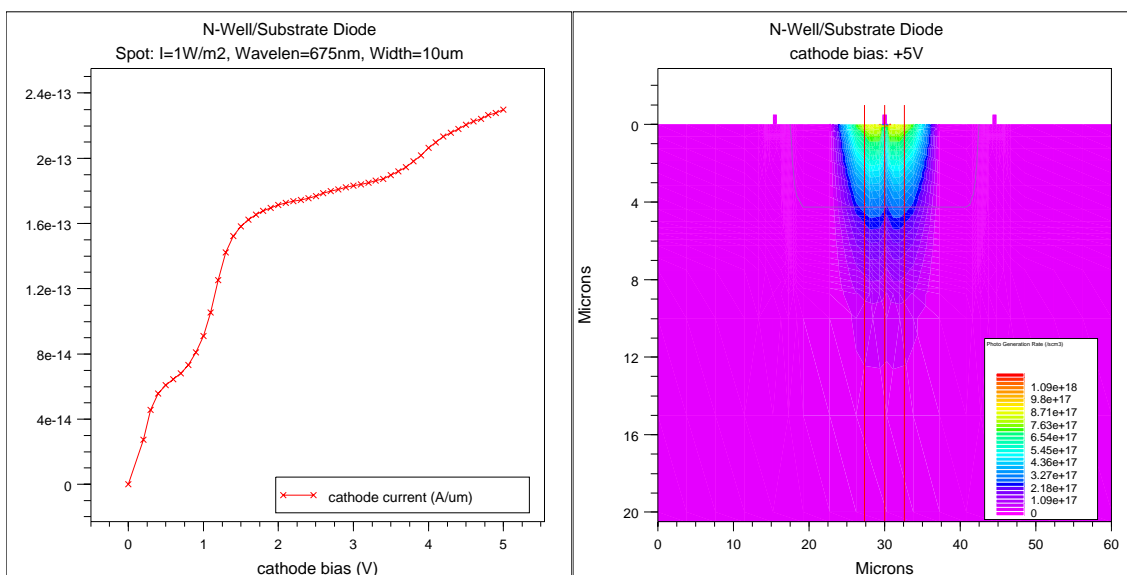


Abbildung 6.2: Ladungsträgerkonzentrationen ohne (links) und mit Beleuchtung

bei dem kein Licht eingestrahlt wird. Geht man jetzt tiefer in das Substrat hinein, so gelangt man über die aufgrund der Sperrspannung verbreiterte Raumladungszone bis in das p-Gebiet des Substrates ( $x \geq 2\mu\text{m}$ ), das hier als Anode dient. Im p-Gebiet sind aber die Elektronen die Minoritäten, weshalb sich deren Konzentration dort ändert, während sich die Löcherdichte (Majoritäten) nicht merklich ändert. Beiden Gebieten gemeinsam ist also, daß sich durch die elektromagnetischen Wellen dieser Intensität lediglich die Konzentration der Minoritäten des jeweiligen Gebietes beeinflussen lassen. Natürlich entstehen zum Beispiel im n-Gebiet ebensoviele Elektronen wie Löcher durch die Photoabsorption, die *relative* Änderung der Dichten macht sich aber nur bei den Minoritäten bemerkbar. Wie in Kapitel 2 beschrieben, ist die Quasi-Fermienergie proportional dem Logarithmus der Ladungsträgerkonzentration. Das Licht erhöht also im Halbleiter die Quasi-Fermienergie der entsprechenden Ladungsträgersorte, wodurch sich deren Dichten erhöhen.

### 6.1.2 N-Wanne/Substrat Diode

In der Abbildung 6.3 ist der Photostrom der N-Wanne/Substrat Diode gezeigt. Im linken Bild ist wiederum der Photostrom in Abhängigkeit der angelegten Sperrspannung aufgetragen. Im rechten Bild der Abbildung ist in farblicher Abstufung die Photo-Generationsrate



**Abbildung 6.3:** Kathodenstrom in Abhängigkeit der Kathodenspannung

aufgezeichnet. Hier kann man die abschirmende Wirkung des Metallkontaktes der Kathode erkennen, die dazu führt, daß die Generationsrate unterhalb des Kontaktes verringert ist.

In der Abbildung 6.4 ist gut zu erkennen, wie lediglich die Minoritäten des entsprechenden Gebietes vom Standard Spot beeinflusst werden. Der Plot reicht tiefer in das Substrat hinein, da die Raumladungszone der N-Wanne/Substrat Diode tiefer im Substrat liegt.

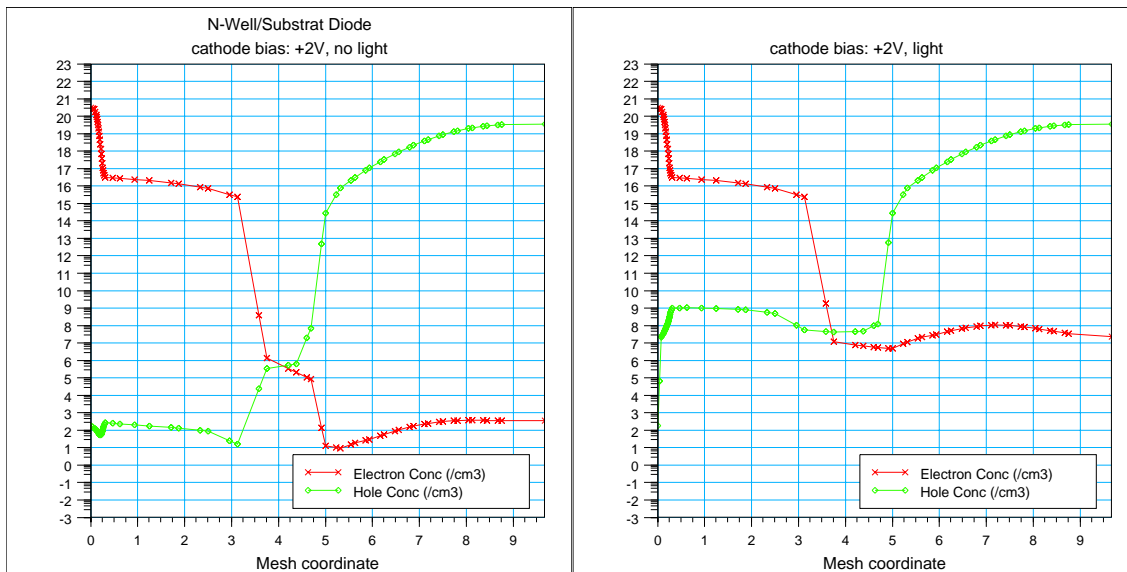


Abbildung 6.4: Ladungsträgerdichten mit (rechts) und ohne Beleuchtung

### 6.1.3 P-Diffusion/N-Wanne Diode

Die P-Diffusion/N-Wanne Diode ist von den dreien die einzige mit einer „echten“ Anode, weswegen hier der Anodenstrom im linken Bild der Abbildung 6.5 aufgetragen ist. Wegen der besseren Vergleichbarkeit ist die Skala der Anodenspannung gespiegelt, so daß wie in den beiden vorherigen Plots die Sperrspannung von links nach rechts ansteigt. Ebenso wurde die Stromachse gespiegelt, damit der Betrag des Photostromes von unten nach oben ansteigt. Im rechten Teil der Abbildung ist wie oben die Photogenerationsrate zu sehen. Sie zeigt ein ähnliches Verhalten, wie die der N-Diffusion/Substrat Diode.

Der Vollständigkeit halber sind auch für diese Diode in der Abbildung 6.6 die Ladungsträgerkonzentrationen ohne und mit Beleuchtung durch den Standard-Spot gezeigt. Hierbei ist im Gegensatz zu den anderen Plots das Potential (rot) mit eingetragen. (Zur Erinnerung: Das Potential  $\Psi = -E_i/q$  ist das *intrinsische* Potential des undotierten Halbleiters; das Fermi-niveau, welches sich durch Dotierung oder durch Einfluß externer Felder einstellt, wird nach Elektronen und Löchern unterschieden und wird als Electron bzw Hole Quasi Fermi Level ( $\Phi_n$  bzw.  $\Phi_p$ ) bezeichnet (siehe auch Kapitel 2).) Das Potential  $\Psi$  folgt den äußeren Spannungen, wie in der Abbildung zu sehen ist. Es ändert sich aber nicht, wenn Licht auf die Diode scheint, denn es ist von außen durch die Kontakte bestimmt. Es werden aber die Quasi-Fermienergien modifiziert, so daß sich die Differenz  $\Psi - \Phi_n$  bzw.  $\Phi_p - \Psi$  erhöht. Dies bewirkt nun wiederum die Erhöhung der verfügbaren Ladungsträger ( $\Psi - \Phi_n \propto \log n$  bzw.  $\Phi_p - \Psi \propto \log p$ ).

Für alle drei Dioden gilt, daß bei geringer Intensität des Lichtes der Photostrom proportional der Intensität des einfallenden Lichtes ist. Die Intensität des Lichtes sei gering, wenn die durch das Licht hervorgerufene Erhöhung der Minoritätenkonzentration die Dichte der Majoritäten im thermischen Gleichgewicht nicht überschreitet.

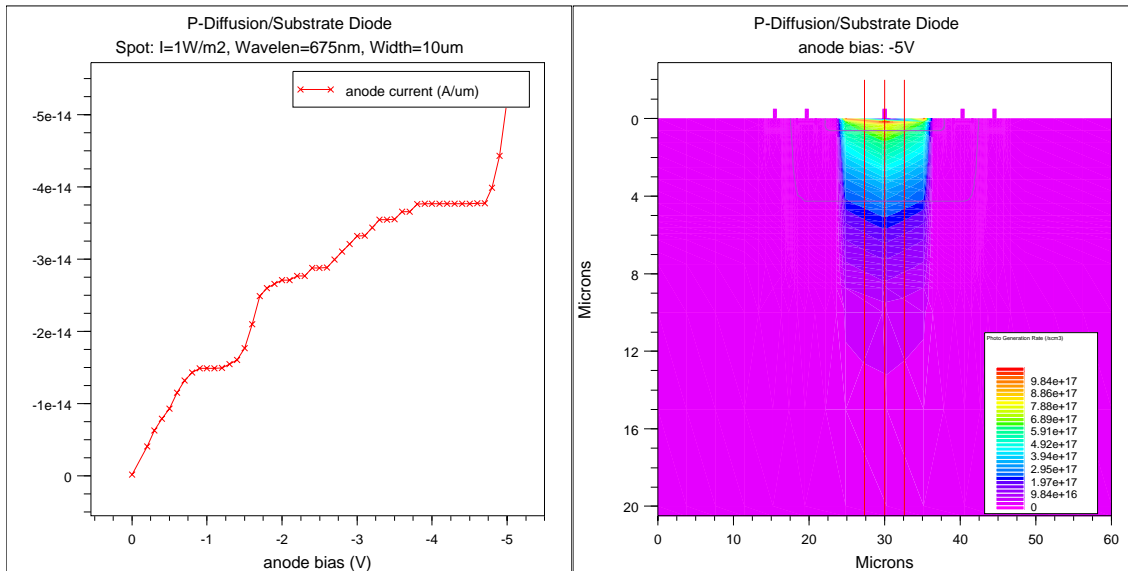


Abbildung 6.5: Anodenstrom in Abhängigkeit der Anodenspannung

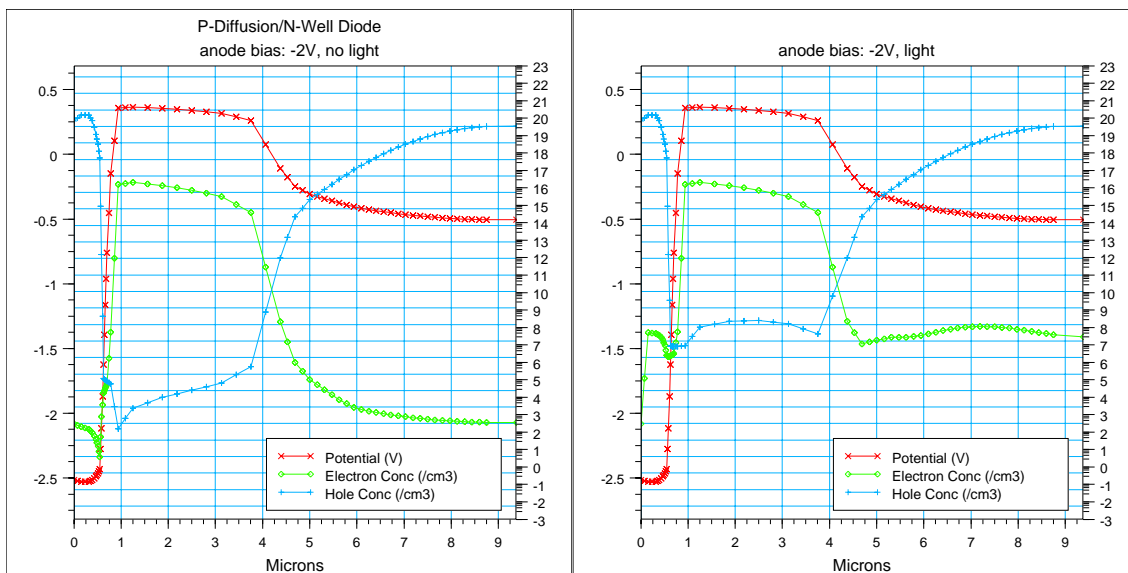


Abbildung 6.6: Ladungsträgerkonzentrationen mit und ohne (links) Beleuchtung

### 6.1.4 Parasitärer Vertikaler PNP-Transistor

Wird ein Transistor als Photo-Sensor benutzt, so wird der Basiskontakt entfernt, da das in die Basis einfallende Licht die Eingangsgröße des Transistors ist. Die Abbildung 6.7 zeigt im linken Teil den Emittierstrom in Abhängigkeit der Emittier-Spannung bei gleichbleibender einfallender Lichtintensität.

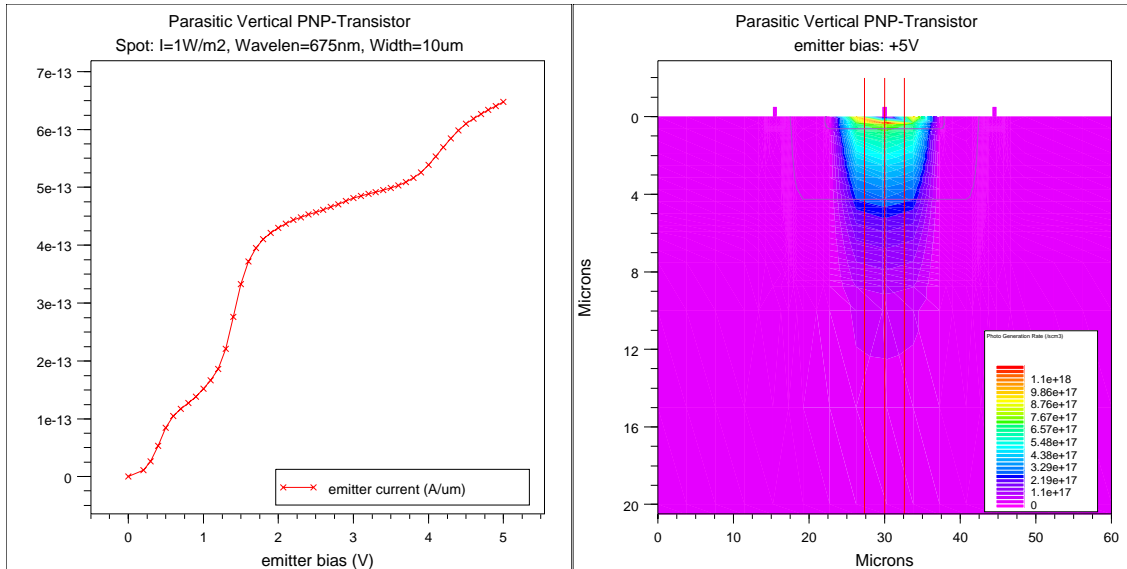


Abbildung 6.7: Emittierstrom in Abhängigkeit der Anodenspannung

Der Verlauf ähnelt dem der N-Wanne/Substrat Diode mit dem Unterschied, daß sich hier die verstärkende Wirkung des Transistors auf den Emittier-Strom auswirkt.

Die Abbildung 6.8 zeigt die Ladungsträgerdichten im Transistor bei Dunkelheit (links) mit einer Emittier-Spannung von +2V und daneben diejenigen mit der Standard-Spot Beleuchtung, ebenfalls mit +2V Emittier-Spannung. Vergleicht man die Plots der P-Diffusion/N-Wanne Diode mit denen des Parasitären Vertikalen PNP-Transistors, so fällt auf, daß im Bereich des Emittier-Basis Übergangs die Ladungsträger dort höher als im thermodynamischen Gleichgewicht sind. Das ist nicht verwunderlich, denn hier wurde im Gegensatz zur P-Diffusion/N-Wanne Diode dem p-Gebiet (beim PNP-Transistor der Emittier) eine positive Spannung beaufschlagt, während bei der P-Diffusion/N-Wanne Diode diesem p-Gebiet (dort Anode) eine negative Spannung auferlegt wurde, um die Diode in Sperrichtung zu betreiben. Weiterhin unterscheiden sich die Wannensubstrat pn-Übergänge. Bei der Diode sind die beiden Potentiale gleich, nämlich 0V, weshalb er sich im thermodynamischen Gleichgewicht befindet. Beim Transistor dagegen hat sich die Raumladungszone des Wannensubstrat Überganges verbreitert. Dazu muß man beachten, daß die N-Wanne (also hier die Basis) *floatet*, das heißt sie hat keinen elektrischen Kontakt nach außen und wird daher nicht mit einem äußeren Potential versehen. Um zu verstehen was in der Basis geschieht dient die Abbildung 6.9. Zunächst sei die rote Kurve betrachtet. Sie stellt das Potential im Transistor ohne Beleuchtung dar. Man erkennt, daß das Potential der Basis ( $0.75\mu\text{m} \leq x \leq 3\mu\text{m}$ ) dem des Emittiers folgt. Das Potential des Emittiers (p-Gebiet) liegt im thermischen Gleichgewicht unter dem



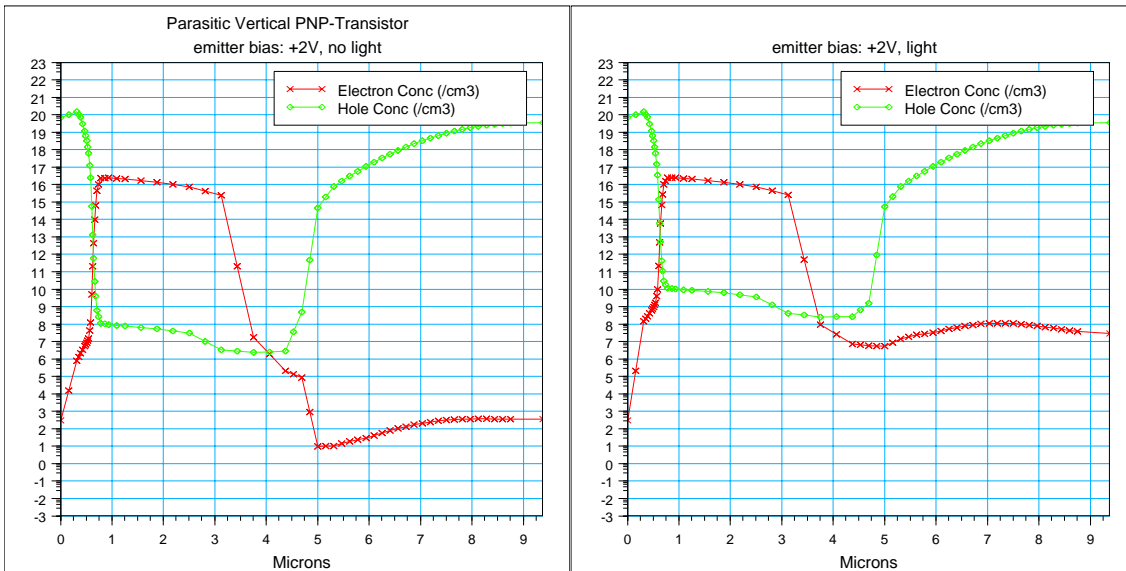


Abbildung 6.8: Konzentration der Ladungsträger mit Beleuchtung (rechts) und ohne

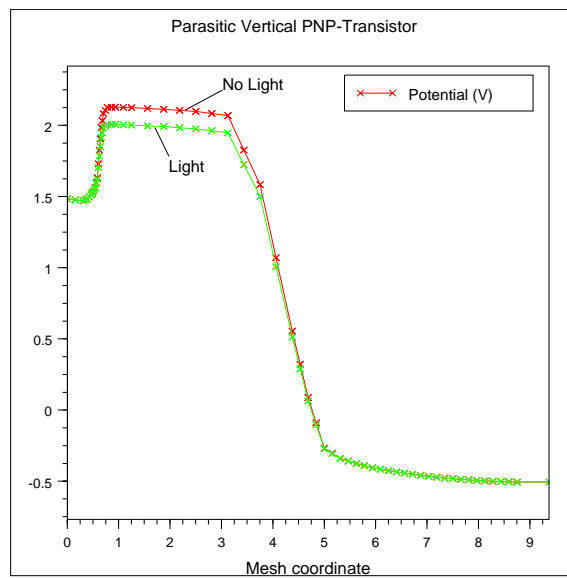


Abbildung 6.9: Vergleich der Potentiale bei Beleuchtung und bei Dunkelheit

Fermi-Niveau, welches dann  $0V$  beträgt. Bei der hier vorliegenden Dotierung beträgt es etwa  $-0.5V$ . Dem Emitter wird aber von außen eine Spannung von  $+2V$  auferlegt, weshalb sein Potential jetzt, wie in der Abbildung zu sehen ist, etwa  $+1.5V$  beträgt. Das Potential, das sich an der dem Emitter anschließenden n-Zone (Basis) einstellt, liegt um die Diffusionsspannung des pn-Überganges (ca.  $0.65V$ ) höher als der Emitter selbst. Das entspräche also einer äußeren Basisspannung von  $+2V$ , wenn die Basis einen Kontakt hätte. Die Basis floatet aber, und wie man sieht folgt ihr Potential dem Emitter-Potential, so daß über der Basis-Emitter Diode keine Spannung abfällt – der Transistor sperrt. Da nun das Basispotential dem Emitterpotential folgt, diese also zusätzlich ein Potential von  $+2V$  erhält, bildet sich zwischen der Basis und dem Kollektor, der auf  $0V$  gelegt ist, eine Raumladungszone aus, da dieser Übergang nun in Sperrichtung geschaltet ist. In der Abbildung 6.8 kann man dies im Gegensatz zur Abbildung 6.6 sehen, in der sich der N-Wannen Substrat Übergang im thermodynamischen Gleichgewicht befindet.

Die grüne Kurve zeigt nun das Potential bei Lichteinfall. Da der Emitter kontaktiert ist und sich die von außen an ihn angelegte Spannung nicht ändert, bleibt sein Potential bei den  $+1,5V$ . Aufgrund der eingestrahelten Photonen und der von ihnen erzeugten Ladungsträgerpaare hat sich nun das Basis-Potential verringert. Das ist also der Fall, der im Gummel Plot (Abbildung 4.7) einer Verringerung des anfänglich auf  $+2V$  eingestellten Basispotentials entspricht. Die Basis-Emitter Diode ist nun in Durchlaßrichtung geschaltet und der Emitter injiziert wegen seines positiveren Potentials Löcher in die Basis. Wie in Kapitel 4 gesehen ist die Basiszone weitgehendst feldfrei und die Löcher diffundieren durch sie hindurch, bis sie schließlich in die Raumladungszone der in Sperrichtung geschalteten Basis-Kollektor Diode gelangen, wo sie aufgrund des negativen Potentials des Kollektors abgesaugt werden. Der Photostrom, der in der Basis erzeugt wird, ist also verstärkt worden. Die Potentialdifferenz in der Basis mit und ohne Licht beträgt aber nur  $\approx 0.12V$ , was einer kleinen Durchlaßspannung entspricht. Wie man ebenfalls in der Abbildung 4.7 erkennen kann, ist die Stromverstärkung  $\beta$  in diesem Bereich nicht sehr groß, so daß die verstärkende Eigenschaft hier nur gering in Erscheinung tritt.

### 6.1.5 BiCMOS NPN-Transistor

Zum Schluß dieses Abschnittes soll noch die Konzentration der Ladungsträger des BiCMOS NPN-Transistors bei Dunkelheit und bei Beleuchtung gezeigt werden. Da der BiCMOS NPN-Transistor wie oben beschrieben einen vom Substrat isolierten Kollektor besitzt, wird dieser mit einer Kollektorspannung von  $+2V$  betrieben. Wie beim PNP-Transistor im vorherigen Abschnitt ist auch hier die Basis ohne Kontakte versehen. In der Abbildung 6.10 kann man wiederum die aufgrund des einfallenden Lichtes, erhöhte Konzentration der Minoritäten erkennen.

Die Abbildung 6.11 zeigt den Vergleich der Potentiale bei Lichteinfall und ohne diesen. Die Basis ist beim NPN-Transistor p-dotiert. Wie man in der Abbildung sieht, erhöht sich das Basispotential nun bei Lichteinfall. Das bedeutet aber, daß die Basis-Emitter Diode wiederum in Durchlaßrichtung geschaltet wird und im Prinzip die gleiche Situation vorliegt wie beim PNP-Transistor, jedoch mit anderen Vorzeichen.

Im nächsten Abschnitt soll untersucht werden, wie die Bauelemente auf einen Lichtstrahl reagieren, wenn dieser an verschiedenen Orten auf den Halbleiter eingestrahlt wird. Dabei

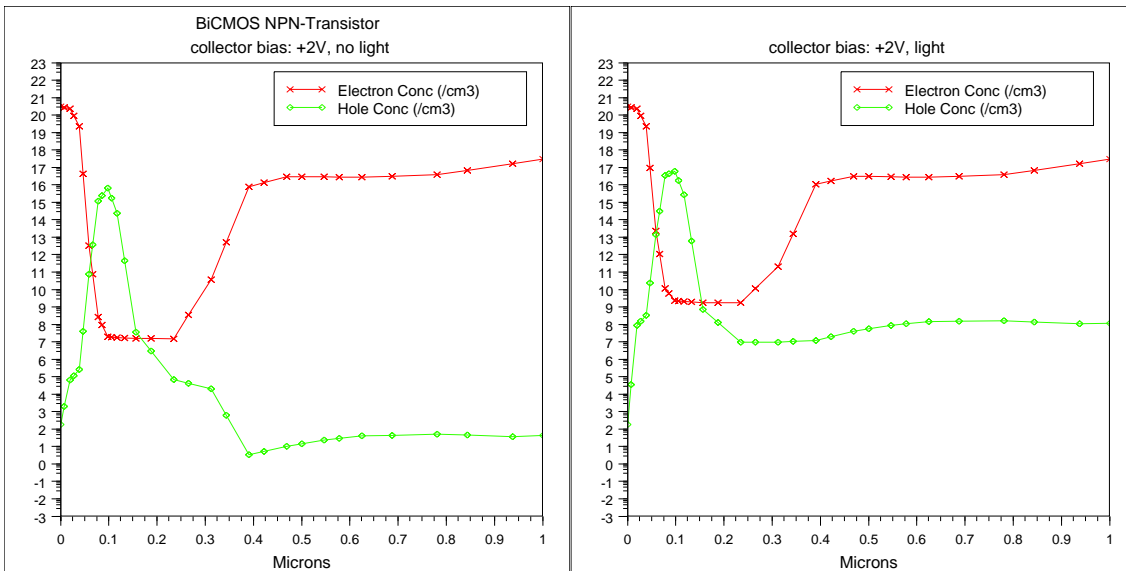


Abbildung 6.10: Ladungsträgerkonzentrationen bei Dunkelheit und bei Beleuchtung

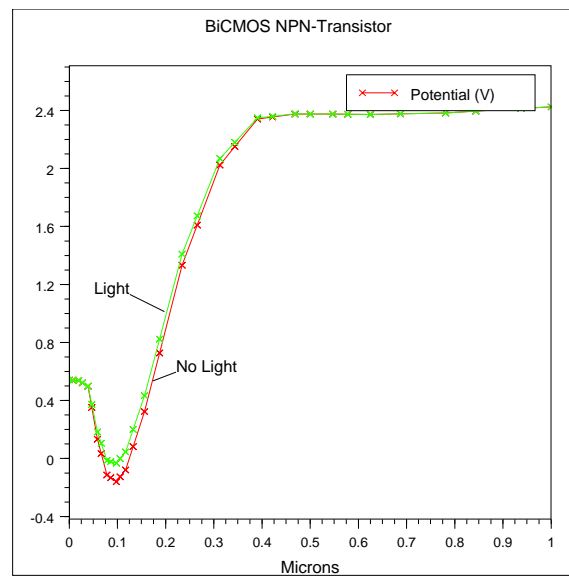


Abbildung 6.11: Ladungsträgerkonzentrationen bei Dunkelheit und bei Beleuchtung

beträgt wie oben die Intensität des Spots  $1W/m^2$ , die Wellenlänge ebenfalls  $675nm$ . Im Unterschied zu oben ist der Spot jedoch nur noch  $1\mu m$  breit.

## 6.2 Räumliche Antwort

In diesem Abschnitt soll untersucht werden, wie sich die Ströme der Bauelemente verhalten, wenn diese an verschiedenen Stellen bestrahlt werden. Um eine einigermaßen akzeptable Auflösung zu erhalten, muß die Breite des Spots erheblich kleiner als die Ausdehnung des entsprechenden Bauelementes sein. Wenn die Intensität des Lichtstrahles gleichbleibt, wie das hier der Fall sein soll, so ist die Verringerung der Breite des Spots mit einer Absenkung der Photogeneration verbunden, da die Gesamtenergie, die auf das Element fällt, kleiner wird. Die Versuche laufen folgendermaßen ab:

Es wird von links nach rechts ein  $1\mu m$  breiter Spot mit einer Intensität von  $1W/m^2$  und einer Wellenlänge von  $675nm$  über das Bauelement bewegt. Die Bauelemente werden also optisch abgescannt und es werden jeweils der Kathoden-, Anoden-, Kollektor-, Emitter- oder Drainstrom gemessen.

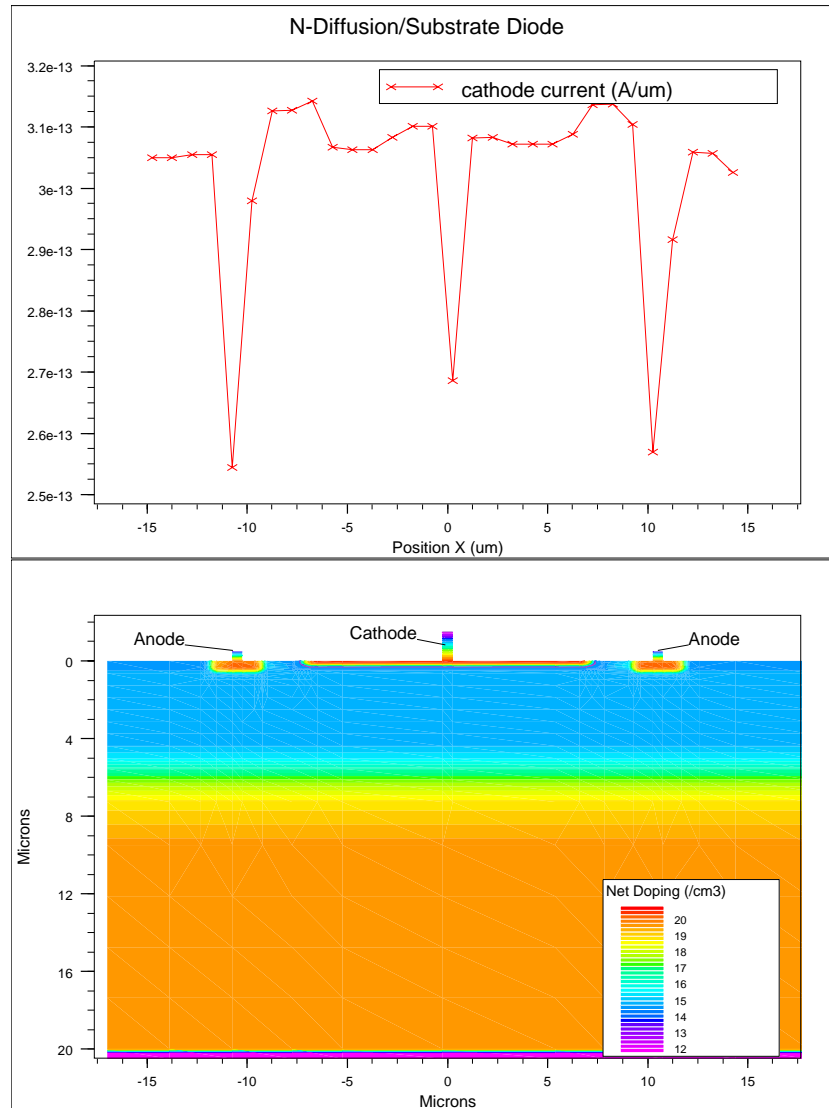
### 6.2.1 N-Diffusion/Substrat Diode

Die Abbildung 6.12 zeigt den Kathodenstrom der N-Diffusion/Substrat Diode in Abhängigkeit von der Position des Lichtspots. Die Kathodenspannung beträgt wieder  $+2V$ , das Substrat ist auf Masse gelegt, die Diode wird also mit einer Sperrspannung von  $2V$  betrieben. Unter der Stromkurve befindet sich zum Vergleich das Layout der Diode. Man kann die deutlichen Einbrüche des Stromes erkennen, wenn der Spot auf die Aluminiumkontakte fällt. Diese schirmen das Licht ab, da die Eindringtiefe von elektromagnetischen Wellen dieser Wellenlänge in Metallen, aufgrund der großen Dichte frei beweglicher Elektronen, äußerst gering ist. Der Lichtspot ist aber größer als das Aluminiumplättchen, so daß er nicht vollständig abgeschirmt wird.

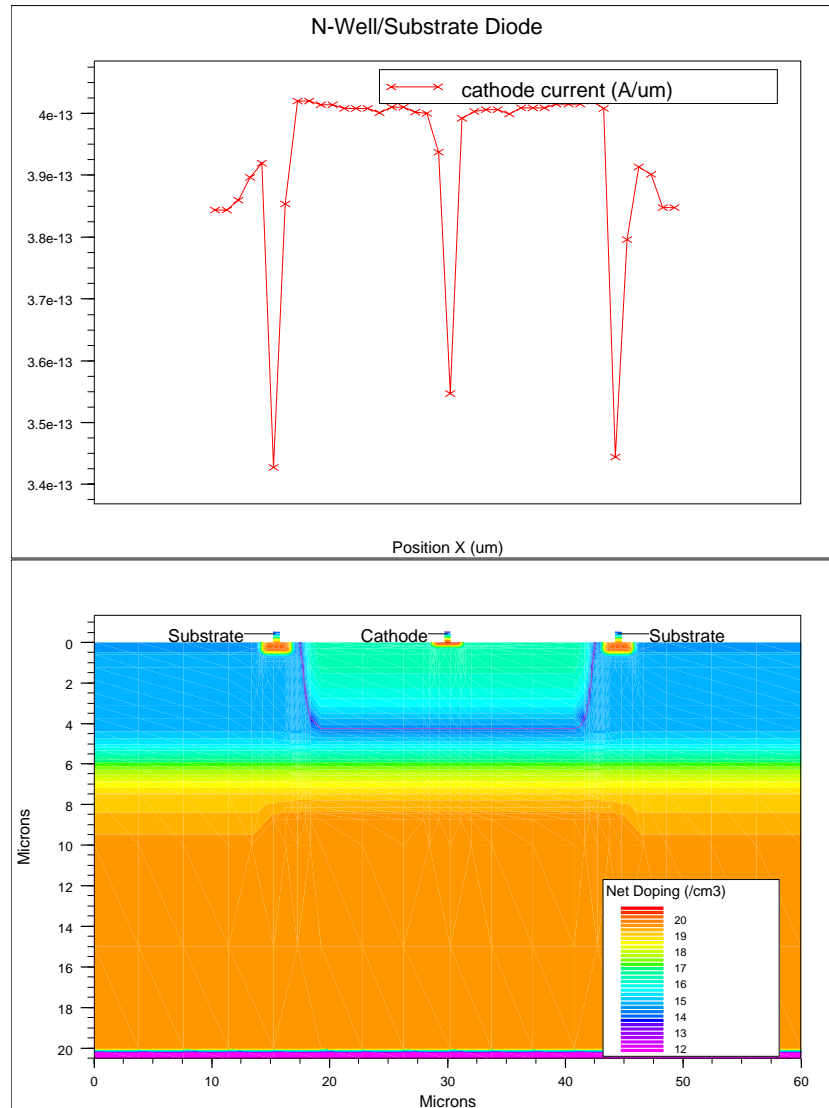
Interessant sind noch die beiden Erhöhungen links und rechts am Anfang bzw. am Ende der n-Diffusion. Dort befinden sich ebenfalls Raumladungszonen, in denen aufgrund des dort herrschenden Feldes die Ladungsträgerpaare abgesaugt werden und zum Photo- bzw. Kathodenstrom beitragen, bevor sie Gelegenheit zur Rekombination haben.

### 6.2.2 N-Wanne/Substrat Diode

In der Abbildung 6.13 ist im oberen Bild die Stromantwort bei räumlicher Variation des Lichtspots der N-Wanne/Substrat Diode zu sehen. Ihre Kathode liegt auf dem externen Potential von  $+2V$ , sodaß die Diode in Sperrichtung betrieben wird. Im unteren Teil der Abbildung ist zum Positionsvergleich des Strahles das Layout der N-Wanne/Substrat Diode gezeigt. Auch hier kann man eine leichte Erhöhung des Kathodenstromes erkennen, wenn das Licht in den Randbereich der Wanne gelangt, wo die erzeugten Ladungsträger in den Einfluß der seitlichen Raumladungszonen gelangen.



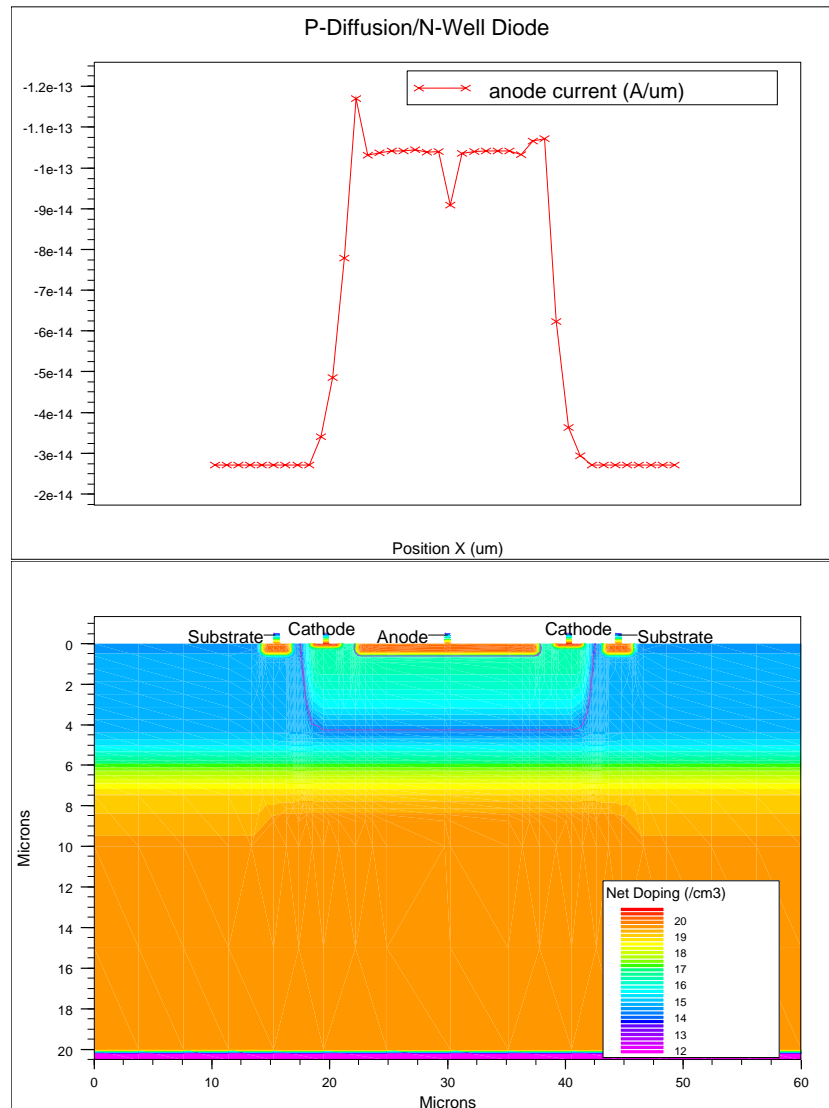
**Abbildung 6.12:** Kathodenstrom einer N-Diffusion/Substrat Diode bei örtlicher Variation des Licht-Spots



**Abbildung 6.13:** Kathodenstrom einer N-Wanne/Substrat Diode bei örtlicher Variation des Licht-Spots

### 6.2.3 P-Diffusion/N-Wanne Diode

Die P-Diffusion/N-Wanne Diode ist wie oben schon erwähnt die einzige der hier besprochenen Dioden, die eine „echte“ Anode besitzt. Diese ist vom Substrat isoliert, und kann sowohl mit positiven als auch mit negativen Potentialen belegt werden. Deshalb wurde hier die Anode auf ein externes Potential von  $-2V$  gebracht, und die Kathode ist auf  $0V$  gelegt. Die Diode wird also wie die anderen mit einer Sperrspannung von  $2V$  betrieben.



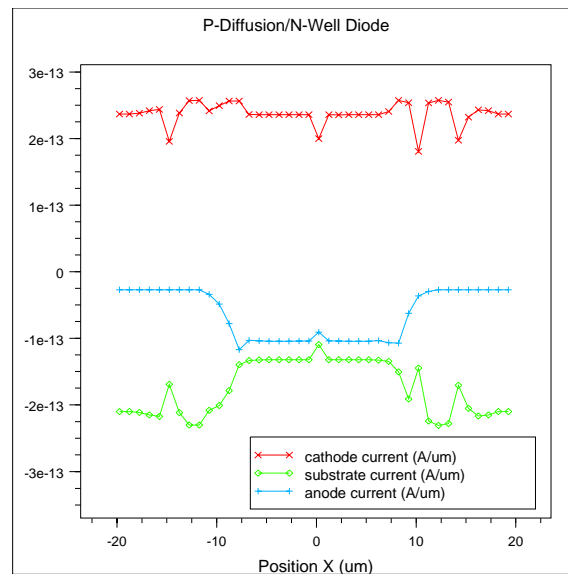
**Abbildung 6.14:** Anodenstrom einer P-Diffusion/N-Wanne Diode bei örtlicher Variation des Licht-Spots

In der Abbildung 6.14 ist daher der Anodenstrom in Abhängigkeit der Position des Lichtspots aufgetragen. Um den Plot mit den anderen besser vergleichen zu können, wurde die Stromachse gespiegelt, so daß der Betrag des Anodenstromes von unten nach oben ansteigt.

Man erkennt deutlich den Einbruch des Anodenstromes, an der Stelle, an der der Spot den Aluminiumkontakt der Anode beleuchtet. Ebenso ist eine Erhöhung des Betrages des Anodenstromes im Bereich der seitlichen Raumladungszonen ( $x \approx 22.5\mu\text{m}$  bzw.  $x \approx 37.5\mu\text{m}$ ) zu beobachten. Die Asymmetrie liegt daran, daß sich der Lichtspot in  $1\mu\text{m}$  Schritten bewegt und nicht exakt spiegelbildlich zur Symmetrie-Achse der Diode verläuft.

Bemerkenswert ist der steile Abfall des Anodenstromes, wenn der Spot in die Nähe der Wanne-Substrat Übergangszone gelangt. Dieser ist so stark, daß die abschirmende Wirkung der beiden Kathodenkontakte, links und rechts am Rande der N-Wanne, sich nicht mehr bemerkbar machen.

Was geschieht nun aber mit den Ladungsträgerpaaren, die auch dann generiert werden, wenn das Licht jenseits der Wanne auf den Halbleiter trifft? Um dies zu erklären, zeigt die Abbil-



**Abbildung 6.15:** Anoden-, Kathoden- und Substratstrom bei örtlicher Variation des Licht-Spots

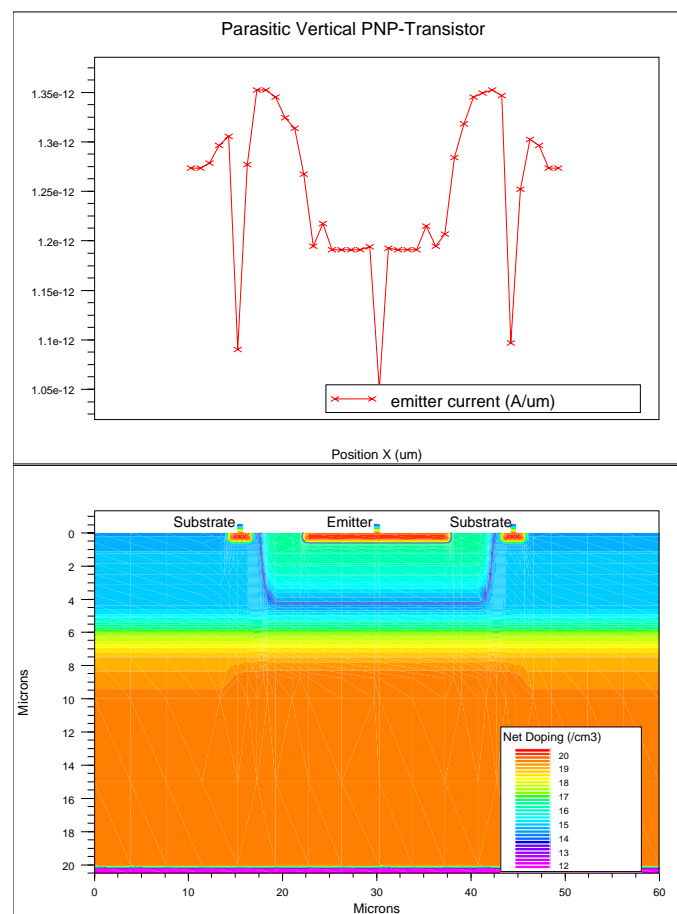
dung 6.15 den Anodenstrom und zusätzlich den Kathoden- und Substratstrom, bei gleichen Lichtverhältnissen wie in Abbildung 6.14. Zu beachten ist hier, daß erstens die Stromachse wieder nach absoluten Größen geordnet ist und zweitens ist die x-Achse um  $30\mu\text{m}$  geschiftet ist, so daß sich die Mitte der Diode bei  $x = 0\mu\text{m}$  befindet. Der Anodenstrom, der dunkelblau eingezeichnet ist, ist derselbe wie in Abbildung 6.14. Zusätzlich sind aber der Kathoden- und der Substratstrom (grün) aufgetragen. Betrachtet man den Kathodenstrom (rot) außerhalb der Wanne, wo der Anodenstrom auf seinen Dunkelwert abgefallen ist, so sieht man, daß dieser, nachdem er den Einbruch über dem Kathodenkontakt erlitten hat, wieder auf den Wert steigt, den er innerhalb der Wanne hatte. Dieser Strom fließt aber nicht über die Anode ab, sondern wird vom Substrat abgeleitet. Bildet man die Summe von Kathoden- und Substratstrom, so erhält man genau den negativen Anodenstrom. Die Raumladungszone des Wannen-Substrat Überganges schirmt also die außerhalb der Wanne erzeugten Ladungsträger von der Anode ab. Das hat den Vorteil, daß zwei benachbarte P-Diffusion/N-Wanne Dioden, die zum Beispiel als jeweils ein Pixel einer optischen Matrix zur Bilderfassung dienen, kein



gegenseitiges Übersprechen zeigen. Um diese Aussage zu verifizieren, stehen Messungen aus, die noch nicht vorgenommen worden sind. In der Simulation jedoch ist der Effekt deutlich zu erkennen, vor allem wenn man den Anodenstrom mit dem Strom der Kathode vergleicht.

#### 6.2.4 Parasitärer Vertikaler PNP-Transistor

In diesem Abschnitt wird wiederum mit dem gleichen Lichtspot wie bei den Dioden der Parasitäre Vertikale PNP-Transistor abgescannt. Dabei ist der Emitter auf  $+2V$  gelegt und der Kollektor (=Substrat) auf  $0V$ . Der obere Teil der Abbildung 6.16 zeigt diesen Plot. Da der Emitterkontakt der einzig zugängliche ist, wird dessen Strom aufgetragen. Das untere Bild zeigt das Layout des Parasitären Vertikalen PNP-Phototransistors. Wird der Transistor als Phototransistor benutzt, so sind wie oben erwähnt keine Basiskontakte vorhanden – die Basis floatet. Deutlich zu sehen ist das Abfallen des Emitterstromes, wenn der Spot auf die



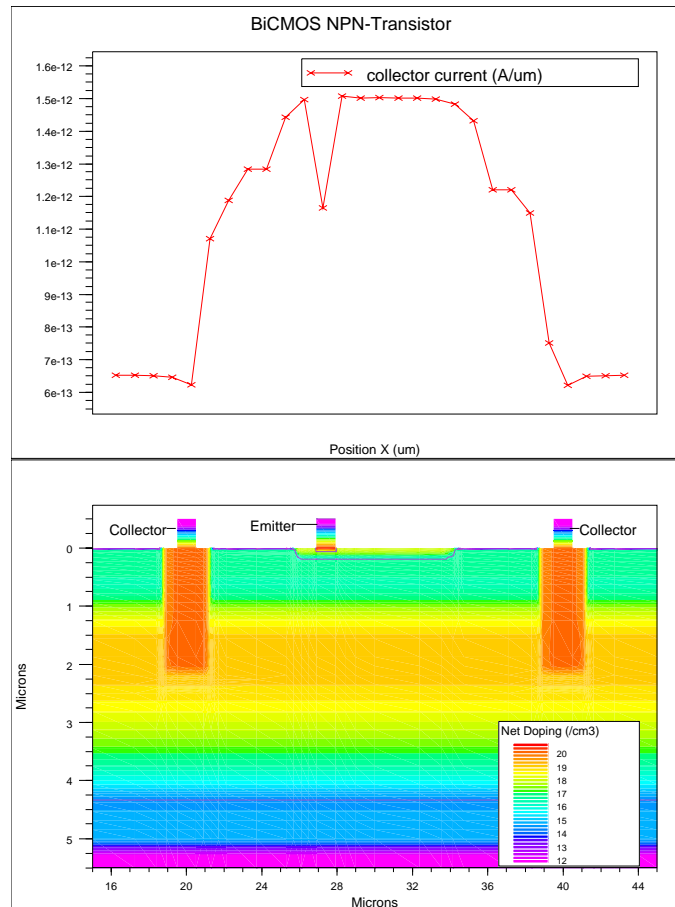
**Abbildung 6.16:** Emitterstrom bei örtlicher Variation des Licht-Spots

Kontakte trifft. Interessant an diesem Plot ist aber das Ansteigen des Emitterstromes, wenn dieser in den Bereich des Wanne-Substrat Überganges (= Basis-Kollektor Übergang) gelangt. Aufgrund des offenen Basiskontaktes folgt das Potential wie in Abbildung 6.9 zu sehen dem des Emitters, wodurch sich eine Basis-Kollektor Sperrspannung aufbaut. Diese Sperrspannung

erzeugt eine Raumladungszone, mit dem von ihr erzeugten elektrischen Feld, welches dann in der Lage ist, die in seiner Nähe generierten Ladungsträger an sich zu ziehen und über die Substratkontakte abzuleiten. Der Strom am Rande der Basis ist, wie in der Abbildung 6.16 zu sehen, höher als in der Nähe des Emitters, was zum Teil auf die größere räumliche Ausdehnung des Feldes der Basis-Kollektor Diode zurückzuführen ist. Außerdem sind wie oben gesehen für die verstärkende Wirkung des Transistors gerade die in der Basis erzeugten und die durch sie hindurchdiffundierenden Ladungsträger verantwortlich. Am Rande der Wanne tragen dann wieder die seitlichen Raumladungszone zur Verringerung des Basispotentials (siehe Abbildung 6.9) bei.

### 6.2.5 BiCMOS NPN-Transistor

Der BiCMOS NPN-Transistor besitzt einen vom Substrat isolierten Kollektor, weshalb dieser nach außen hin verfügbar ist und deshalb hier mit einem Potential von  $+2V$  belegt wird. Der Emitter ist auf  $0V$  gelegt, und die Basis hat wie beim PNP-Phototransistor keinen Kontakt nach außen. In der Abbildung 6.17 ist der Kollektorstrom in Abhängigkeit der Spotposition



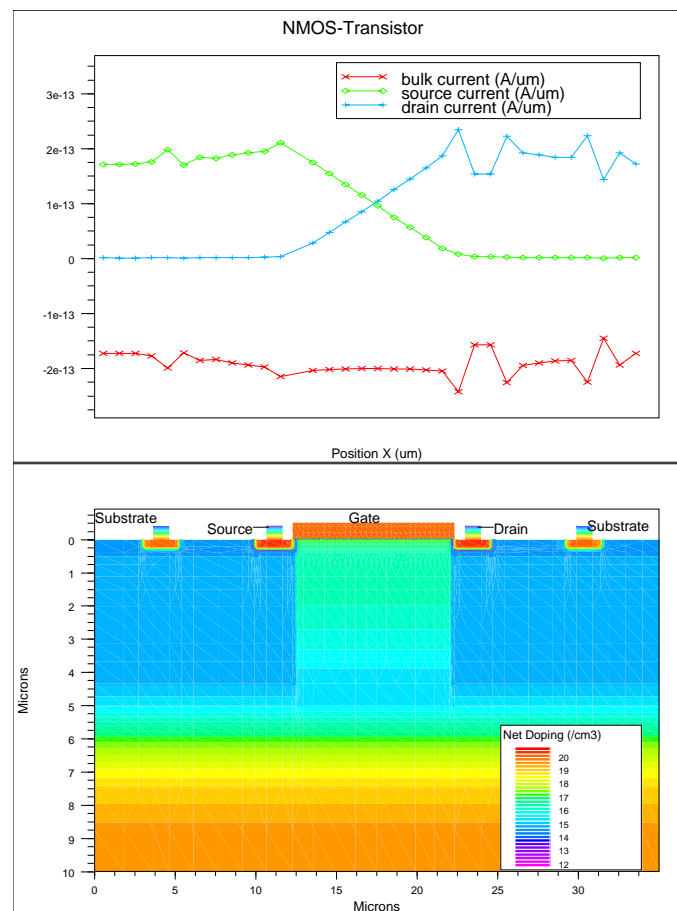
**Abbildung 6.17:** Kollektorstrom bei örtlicher Variation des Licht-Spots

aufgetragen. Aufgrund der Isolierung des Bauteiles vom Substrat ist der Kollektorstrom rela-

tiv gut abgegrenzt. Lediglich ein schwaches Anwachsen links und rechts der Kollektorkontakte ist zu beobachten. Dies läßt sich wohl auf den Kollektor- (bzw. Sink-) Substrat Übergang zurückführen, der durch das Kollektorpotential von  $+2V$  in Sperrichtung geschaltet ist und so als „Photodiode“ zum Kollektorstrom jenseits des Transistors beiträgt. Der Strom weist hier im Maximum den größten Wert aller Bauteile auf, was an der relativ hohen Stromverstärkung  $\beta$  liegt.

### 6.2.6 NMOS-Transistor

Zum Abschluß soll noch das Verhalten des NMOS-Transistors beim optischen Scannen untersucht werden. In der Abbildung 6.18 sind der Drainstrom (blau), der Source- und der Sub-



**Abbildung 6.18:** Drain-, Source- und Substratstrom bei örtlicher Variation des Licht-Spots

stratstrom (rot) eingetragen. Wenn mit *Substrat* die Umgebung<sup>1</sup> eines Bauelementes gemeint ist, wird dieses auch manchmal als *Bulk* (engl.: Masse, Hauptteil) bezeichnet. Im oberen Bild wird der Substratstrom daher als „bulk current“ bezeichnet. Der NMOS-Transistor wird

<sup>1</sup>Beim PMOS-Transistor wäre dies die n-Wanne

nicht extra als Phototransistor präpariert bzw. auch nicht als solcher eingesetzt. Hier sollen aber dennoch seine optischen Eigenschaften untersucht werden, denn er wird, wenn auch ungewollt, vom einfallenden Licht beeinflusst. Außer dem Drainpotential, das auf  $+3V$  liegt, betragen alle anderen Potentiale (Source, Gate, Substrat) des Transistors  $0V$ .

Von links kommend ändert sich der Drainstrom erst dann, wenn sich der Spot in der Nähe bzw. innerhalb der Kanalregion befindet. Davor fließen die vom Spot erzeugten Ladungsträger über Source und Substrat ab. In der Kanalregion bleibt der Substratstrom relativ konstant und ein Teil der Ladungsträger fließt schon über den Drain-Kontakt ab. In der Mitte des Kanals teilen sich Source und Drain den Strom. Je näher der Spot nun zum Drain-Gebiet wandert, desto höher wird der Drainstrom, wobei der Source-Strom abfällt. Jenseits des Kanals fällt der Spot nun in die Raumladungszone der in Sperrichtung gepolten Drain-Substrat Diode, weswegen der Strom kurz ansteigt, dann aber wieder abfällt, weil der Drainkontakt das Licht abschirmt. Der Source-Strom ist nun auf Null abgefallen, der Transistor ist an den Photoströmen nur noch durch die Drain-Substrat Diode beteiligt. Der Transistor ist vom Layout her gesehen spiegelsymmetrisch. Das unterschiedliche Verhalten der Source- und Drainströme ist also ausschließlich ein Effekt der äußeren Potentiale.

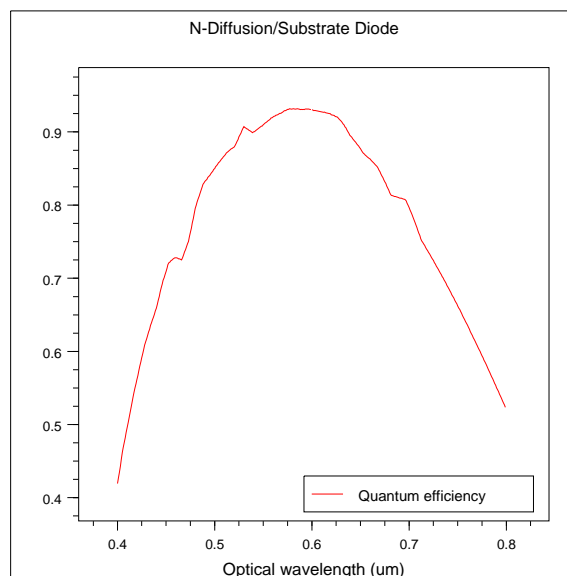
Es fließt an keiner Position des Lichtspots ein Strom von Drain über den Kanal nach Source. Wäre dies der Fall, so müßte der Source-Strom negativ sein. Hier kommen aber nur positive Ströme an Source und Drain vor, was bedeutet, daß immer nur ein Strom von Source in das Substrat oder von Drain in das Substrat fließt. Das Licht hat hier also nicht den gleichen Effekt wie bei den Bipolar-Transistoren. Dort verursachte der Einfall des Lichtes in die Basiszone ebenso einen Kollektorstrom, wie ein entsprechendes von außen angelegtes Basispotential. Das ist beim NMOS nicht der Fall. Eine positive Gate-Spannung läßt einen Strom von Drain nach Source fließen, sofern eine positive Drain-Source Spannung anliegt. Das Licht jedoch läßt nur Ströme über einerseits die Source-Substrat Diode oder andererseits über die Drain-Substrat Diode fließen.

## 6.3 Spektrale Antwort

In diesem Abschnitt soll das spektrale Verhalten der Bauelemente betrachtet werden. Das heißt, ein Lichtspot mit unterschiedlicher Wellenlänge fällt auf das Bauteil und die entsprechenden Ströme werden erfaßt. Es werden Bauteile mit einem Spot beleuchtet, der eine Intensität von  $1W/m^2$  aufweist und  $10\mu m$  breit ist. Im Unterschied zu Abschnitt 6.1 wird jetzt aber jeweils die Wellenlänge von  $0.4\mu m$  bis  $0.8\mu m$  variiert. Gemessen (bzw. vom Simulator berechnet) wird dann der entsprechende Strom des beleuchteten Bauelementes in Abhängigkeit der Wellenlänge des eingestrahnten Lichts. Damit die verschiedenen Plots miteinander vergleichbar sind, wird der Strom auf den Photo-Strom der Lichtquelle normiert. Damit erhält man dann eine Quanteneffizienz, die das Verhältnis des im Bauteil hervorgerufenen Stromes zum Photostrom der Quelle ausdrückt.

### 6.3.1 N-Diffusion/Substrat Diode

Die Diode unterliegt einer Sperrspannung von  $2V$ . Die Abbildung 6.19 zeigt die vorher definierte Quanteneffizienz in Abhängigkeit der Wellenlänge des einfallenden Lichtes.



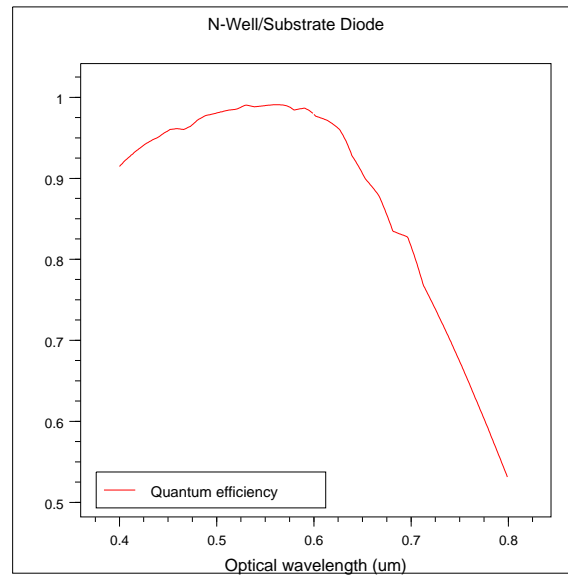
**Abbildung 6.19:** Verhältnis des Kathodenstromes zum einfallenden Photostrom in Abhängigkeit der Wellenlänge des einfallenden Lichtes

Das Maximum der Kurve liegt bei ca.  $580\text{nm}$ . Das Abfallen nach größeren Wellenlängen hin wird durch die hohe Eindringtiefe der Photonen in den Siliziumkristall bestimmt. Sie erzeugen Ladungsträgerpaare tief im Substrat, die dann, bevor sie zur Raumladungszone diffundieren, rekombinieren und nicht zum Strom beitragen. Auf der kurzwelligen Seite des Maximums besitzen die Photonen eine sehr geringe Eindringtiefe. Sie werden in der hoch dotierten n-Diffusion absorbiert. Aufgrund der hohen Dotierung herrschen dort aber sehr viele Rekombinationszentren, so daß diese ebenfalls keinen Beitrag zum Kathodenstrom leisten.

### 6.3.2 N-Wanne/Substrat Diode

Die Kathode der N-Wanne/Substrat Diode ist ebenfalls auf ein Potential von  $+2\text{V}$  gelegt. Die Abbildung 6.20 zeigt ihre Quantenausbeute.

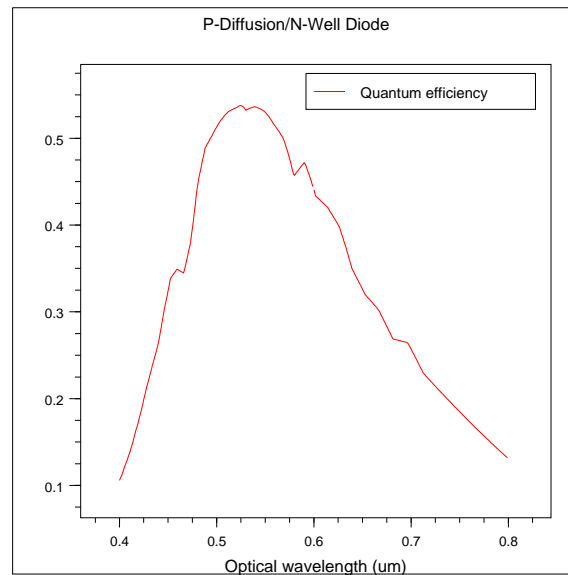
Hier liegt das Maximum bei ca.  $560\text{nm}$ . Sie hat im langwelligen Bereich eine gering höhere Quanteneffizienz. Aufgrund der tiefer im Substrat liegenden pn-Übergangszone würde man erwarten, daß diese Diode im langwelligen spektralen Bereich mehr Photonen absorbieren kann als die N-Diffusion/Substrat Diode. Der Effekt ist aber nur gering. Eine deutliche Änderung ist aber im kurzwelligeren Bereich zu erkennen. Die energiereicheren Photonen, die nur eine geringe Eindringtiefe haben, erzeugen nun Ladungsträger in einem relativ schwach dotiertem Gebiet. Das ist das Gebiet der n-Wanne. Im Vergleich zur n-Diffusion ist sie um ca. 5 Größenordnungen schwächer dotiert. Das heißt also die erzeugten Ladungsträger können relativ weit diffundieren, ehe sie eine Möglichkeit zur Rekombination haben. Sie gelangen so in den Bereich des Wannens-Substrat Überganges und tragen zum Kathodenstrom bei.



**Abbildung 6.20:** Verhältnis des Kathodenstromes zum einfallenden Photostrom in Abhängigkeit der Wellenlänge des einfallenden Lichtes

### 6.3.3 P-Diffusion/N-Wanne Diode

Bei der P-Diffusion/N-Wanne Diode wird die Anode mit einem Potential von  $-2V$  belegt. Die Abbildung 6.21 zeigt daher das Verhältnis des Anodenstromes zum Photostrom der Lichtquelle.

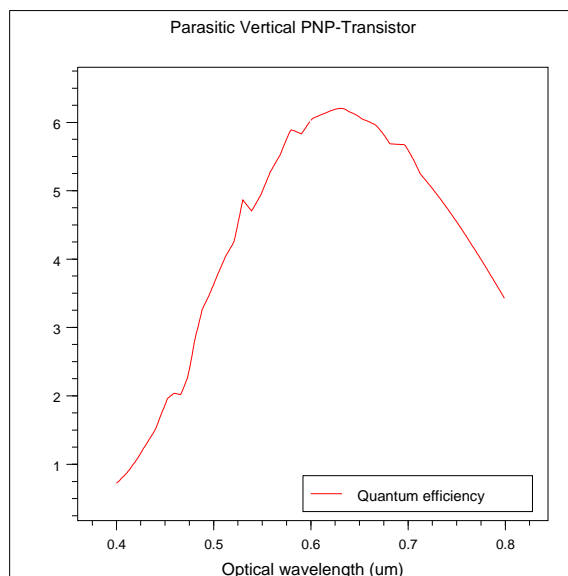


**Abbildung 6.21:** Verhältnis des Anodenstromes zum einfallenden Photostrom in Abhängigkeit von der Wellenlänge des einfallenden Lichtes

Die Quanteneffizienz dieser Diode hat ihren maximalen Wert bei ca.  $530nm$ . Der qualitative Verlauf ist dem der N-Diffusion/Substrat Diode ähnlich. Das liegt daran, daß auch bei der P-Diffusion/N-Wanne Diode das p-Gebiet der Anode stark dotiert ist, so daß die kurzwelligen Photonen, die nur eine geringe Eindringtiefe besitzen, aufgrund der hohen Dotierung in der Anode rekombinieren, bevor sie in den Einfluß des Feldes der Raumladungszone gelangen.

### 6.3.4 Parasitärer Vertikaler PNP-Transistor

Wie bei den anderen optischen Simulationen wurden hierfür die Basiskontakte des Transistors entfernt. Der Emitter liegt auf einem externen Potential von  $+2V$ , und die Basis floatet. Die Abbildung 6.22 zeigt die Quanteneffizienz des PNP-Transistors.



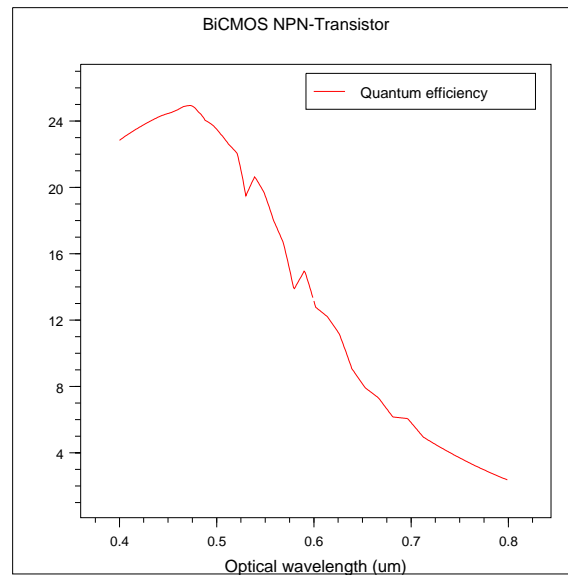
**Abbildung 6.22:** Verhältnis des Emitterstromes zum einfallenden Photostrom in Abhängigkeit der Wellenlänge des einfallenden Lichtes

Es fällt zunächst auf, daß die Quanteneffizienz höher als eins ist. Das kommt daher, daß die Ladungsträger in der Basis erzeugt werden, und dann durch den Transistor verstärkt werden. Um also die echte Umwandlung in den von den Photonen im Bauteil erzeugten Strom zu erhalten, müßte der Photostrom noch durch die Stromverstärkung  $\beta$  geteilt werden. Benutzt man allerdings einen Phototransistor, dann will man gerade die verstärkende Eigenschaft dieses Bauelementes nutzen, so daß hier darauf verzichtet wurde die Verstärkung wieder herauszurechnen. Das Maximum der Kurve liegt bei ca.  $630nm$ . Zum kurzwelligen Ende hin gilt qualitativ das was bei der P-Diffusion/N-Wanne oder N-Diffusion/Substrat Diode gesagt wurde. Die kurzwelligen Photonen mit geringer Eindringtiefe finden auch hier eine hochdotierte Zone (Emitter) vor, die ihnen viele Rekombinationsmöglichkeiten liefert, weshalb sie keinen Beitrag zum Strom leisten können. Am langwelligen Ende besitzt der PNP-Transistor eine relativ hohe Quanteneffizienz. Das liegt einerseits daran, daß die Stromverstärkung  $\beta$  keine Konstante ist. Sie ist nämlich eine Funktion der Basis-Emitter Spannung. Da die Photonen mit unterschiedlichen Wellenlängen unterschiedliche Energie besitzen, erzeugen sie auch

unterschiedliche Basis-Emitter Potentialdifferenzen. Außerdem ist durch die floatende Basis die Basis-Kollektor Sperrschicht nicht konstant und führt zu einer Variation des elektrischen Feldes in der Raumladungszone. Je nach Größe und räumlicher Ausbreitung gelingt es diesem Übergang mehr oder weniger gut, Ladungsträger abzusaugen.

### 6.3.5 BiCMOS NPN-Transistor

Der BiCMOS NPN-Transistor hat einen vom Substrat isolierten Kollektor. Daher wird bei ihm dieser auf  $+2V$  hochgefahren, wobei Emitter und Substrat auf  $0V$  liegen. Als Phototransistor floatet die Basis wieder.



**Abbildung 6.23:** Verhältnis des Kollektorstromes zum einfallenden Photostrom in Abhängigkeit der Wellenlänge des einfallenden Lichtes

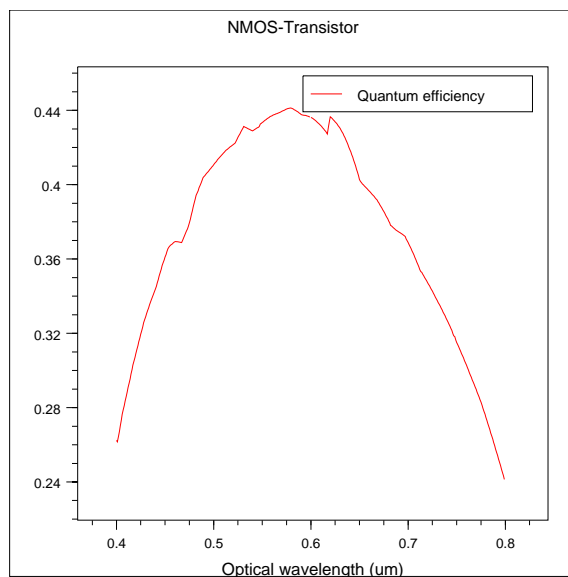
Die Abbildung 6.23 zeigt das Verhältnis des Kollektorstromes zum einfallenden Photostrom der Lichtquelle. Was die Höhe der Quanteneffizienz betrifft, so gilt das wie beim PNP-Transistor gesagten. Der von den Photonen erzeugte Strom in der Basis wird vom Transistor verstärkt, weshalb die Quanteneffizienz größer als eins ist. Sie liegt aufgrund der größeren Stromverstärkung des BiCMOS NPN-Transistors ebenfalls höher als die des PNP. Das Maximum liegt bei ca.  $470nm$ . Der BiCMOS NPN-Transistor besitzt eine Basis, die sich sehr nahe an der Oberfläche befindet. Weil diese nicht so stark dotiert ist (und auch nicht sein darf), rekombinieren die von den kurzwelligen Photonen erzeugten Ladungsträgerpaare nicht so schnell, so daß sie durch die Basis hindurchdiffundieren können, zum Kollektor gelangen, und somit zum Strom beitragen. Aufgrund dieser Beschaffenheit der Basis und ihrer Nähe zur Oberfläche hat der Transistor sein Maximum in der Quanteneffizienzkurve bei relativ kleinen Wellenlängen. Der PNP-Transistor hat wohl auch eine schwach dotierte Basis, aber vor der Basis kommt noch die hoch dotierte Emitter-Zone nahe der Oberfläche, die viele Rekombinationsmöglichkeiten bietet. Im langwelligen Bereich schneidet der BiCMOS NPN-Transistor relativ schlecht ab, da diese Photonen eine große Eindringtiefe haben und zu weit von der



Basis entfernt sind, als daß sie in ihren Wirkungsbereich gelangen. Außerdem befindet sich in einer Tiefe von ca.  $2\mu\text{m}$  der hoch dotierte N-Burried Layer, in dem die Ladungsträgerpaare viele Rekombinationszentren vorfinden und dann dem Strom verloren gehen.

### 6.3.6 NMOS Transistor

Der NMOS-Transistor wird nicht explizit als Phototransistor ausgelegt, weswegen er alle ursprünglichen Kontakte besitzt. Außerdem war oben zu sehen, daß er bei Lichteinfall gar nicht als solcher funktioniert. In der Abbildung 6.24 ist das Verhältnis des Drainstromes zum einfallenden Photostrom gezeigt.



**Abbildung 6.24:** Verhältnis des Drainstromes zum einfallenden Photostrom in Abhängigkeit der Wellenlänge des einfallenden Lichtes

Im Abschnitt 6.2 wurde gezeigt, daß das Licht, welches auf den Transistor fällt, lediglich die pn-Übergänge beeinflusst, und nicht wie bei den Bipolar Transistoren die steuernde Wirkung einer Basisspannung bzw. Gatespannung übernimmt. Es ist daher nicht verwunderlich, daß die Quanteneffizienzkurve im Verlauf derjenigen der N-Diffusion/Substrat Diode ähnelt, da das Drain-Gebiet (ebenso wie das Source-Gebiet) gerade aus einer n-Diffusion besteht. Das Maximum der Kurve liegt wie bei der N-Diffusion/Substrat Diode bei  $580\text{nm}$ . Der Wert der Quanteneffizienz liegt aber nur bei ca. 0.44. Wie aber ebenfalls schon gesehen teilen sich die Photoströme in den Drain- und Sourcestrom auf, so daß der verringerte Wert gegenüber der N-Diffusion/Substrat Diode plausibel wird.

Zu den absoluten Größen aller Quanteneffizienzen, ist zu sagen, daß diese, verglichen mit Messungen, zu hoch ausfallen. Dabei muß berücksichtigt werden, daß hier keine Passivierungsschichten simuliert wurden, die auf realen Bauelementen vorhanden sind, und einen Teil des einfallenden Lichtes reflektieren und absorbieren. Aussagekräftiger ist der Vergleich der hier simulierten Bauelemente untereinander.

## 6.4 Transiente Analyse

Im letzten Abschnitt dieses Kapitels sollen für die Bauelemente noch transiente Analysen gemacht werden. Dabei soll untersucht werden, wie sich die Ströme der Bauelemente verhalten, wenn das Licht zu- oder abgeschaltet wird. Zur transienten Analyse wird der Standard Lichtspot verwendet. Er hat eine Intensität von  $1W/m^2$ , die Wellenlänge beträgt  $675nm$  und seine Breite  $10\mu m$ . Dieser Spot wird zur Zeit  $t = 0$  eingeschaltet. Er bleibt so lange eingeschaltet, bis sich der Strom des entsprechenden Bauelementes nicht mehr ändert. Danach wird das Licht ausgeschaltet und der Stromabfall in Abhängigkeit der Zeit wird vom Simulator aufgezeichnet.

### 6.4.1 N-Diffusion/Substrate Diode

Die Abbildung 6.25 zeigt das zeitliche Verhalten der N-Diffusion/Substrate Diode. Um die

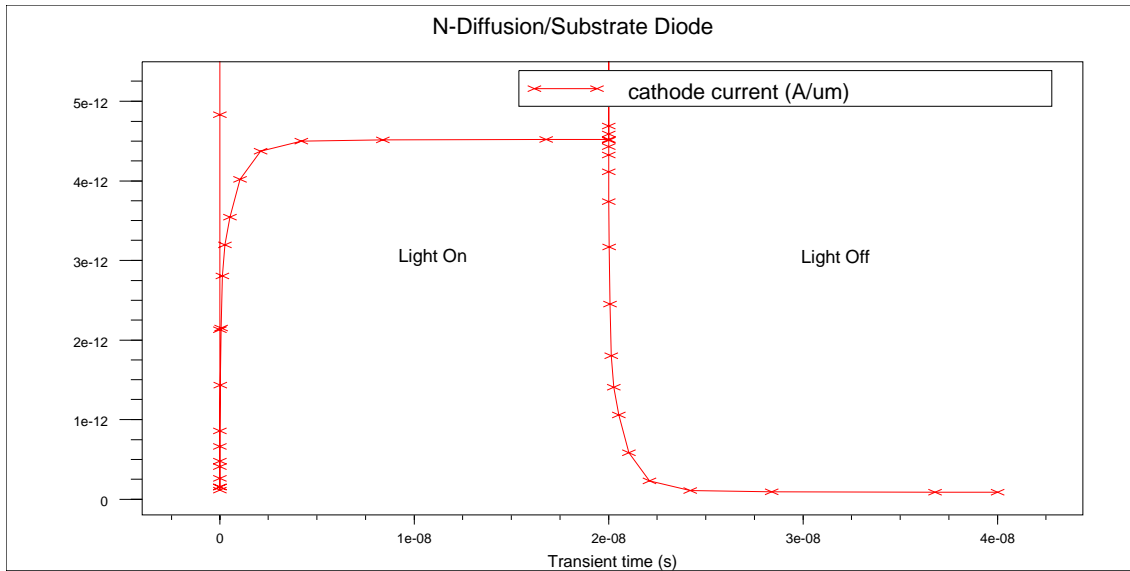


Abbildung 6.25: Kathodenstrom in Abhängigkeit der Zeit

Schaltzeiten der verschiedenen Bauelemente miteinander zu vergleichen, wird die Zeit gemessen, die das Bauteil nach Abschalten des Lichtes benötigt, um von dem Stromwert mit Beleuchtung auf 1/10 dieses Wertes abzufallen. Der Kathodenstrom der N-Diffusion/Substrat Diode fällt in ca.  $1.5ns$  von seinem Wert bei Beleuchtung auf ein Zehntel ab.

### 6.4.2 N-Wanne/Substrat Diode

In der Abbildung 6.26 ist die Antwort der N-Wanne/Substrat Diode auf den Lichtsprung zu sehen. Diese Diode braucht etwa  $3.7ns$  um vom Strom bei Beleuchtung auf dessen Zehntel abzufallen. Die Kathode der N-Wanne/Substrat Diode ist viel schwächer dotiert, als die Kathode der N-Diffusion/Substrat Diode, weshalb die Ladungsträger in der Wanne nicht

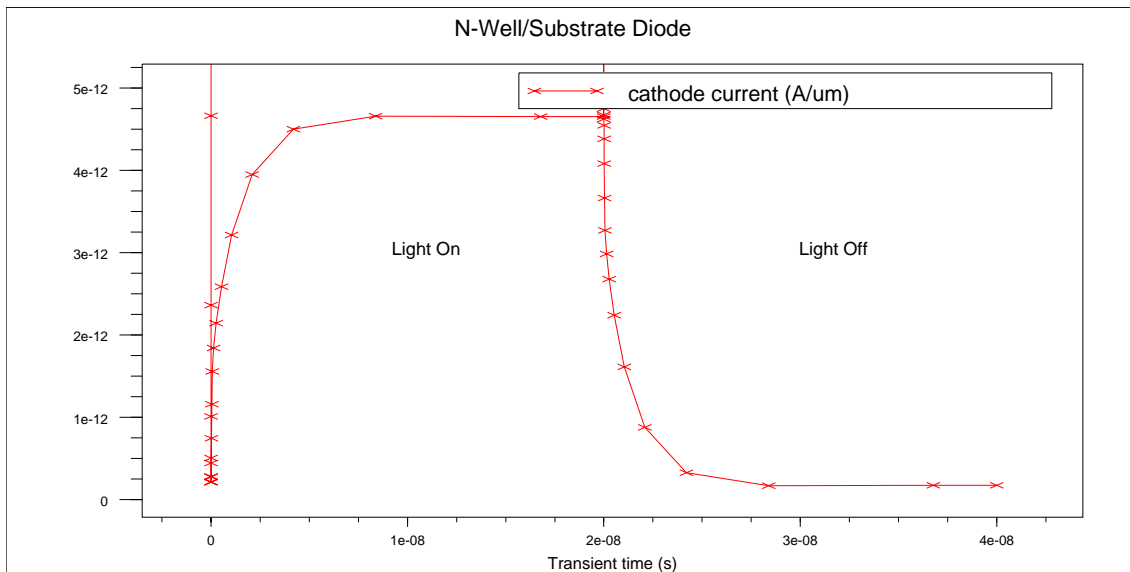


Abbildung 6.26: Kathodenstrom in Abhängigkeit der Zeit

so viele Rekombinationszentren vorfinden und daher eine größere Zeit benötigen, um sich abzubauen.

### 6.4.3 P-Diffusion/N-Wanne Diode

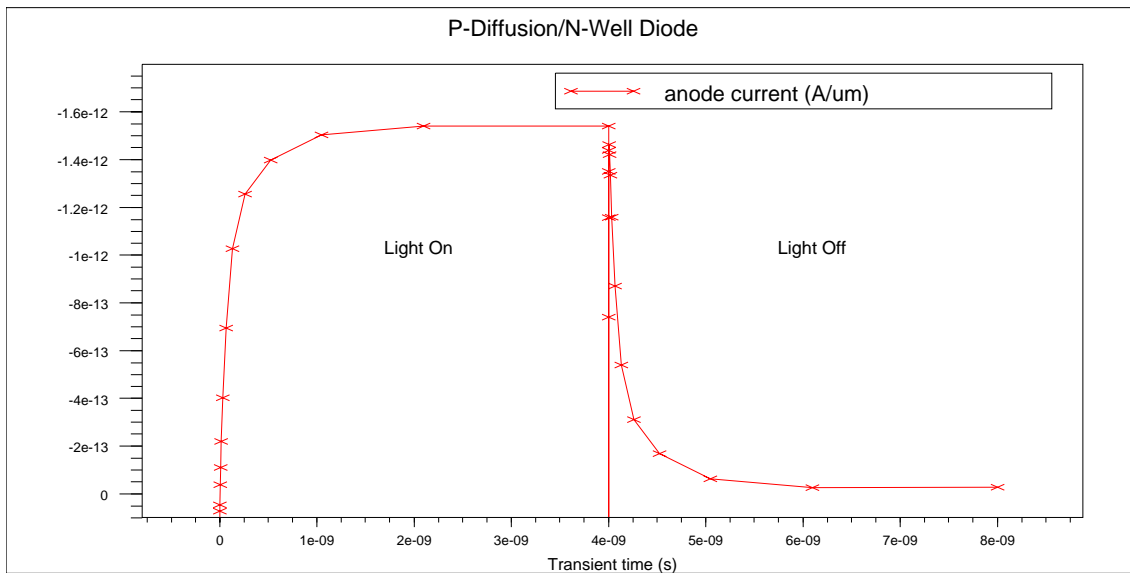


Abbildung 6.27: Anodenstrom in Abhängigkeit der Zeit

Bei dieser Diode ist der Anodenstrom aufgetragen. Da Photodioden in Sperrrichtung betrieben werden, beträgt die Anodenspannung  $-2V$ . Daher ist der Strom negativ. Um aber einen ähn-

lichen Plot wie vorher zu bekommen, wurde in der Abbildung 6.27 die Stromachse gespiegelt, so daß der Betrag des Anodenstromes von unten nach oben ansteigt. Die P-Diffusion/N-Wanne ist die schnellste der hier besprochenen Dioden. Sie benötigt lediglich ca.  $620ps$  um vom Anodenstrom, der bei Beleuchtung fließt, auf  $1/10$  dieses Wertes abzufallen. Die P-Diffusion/N-Wanne ist diejenige Diode mit den höchstdotierten n- und p-Gebieten. Zwar ist die Störstellendichte der n-Diffusion ebenso hoch wie die der p-Diffusion, im Gegensatz aber zur N-Diffusion/Substrat Diode, deren p-Gebiet das Substrat ist, welches relativ schwach dotiert ist, dient bei der P-Diffusion/N-Wanne Diode die Wanne als n-Gebiet, die eine höhere Dotierkonzentration als das Substrat besitzt. Die P-Diffusion/N-Wanne Diode hat von allen Dioden die meisten Rekombinationszentren, so daß ihre Schaltzeiten die kleinsten sind.

#### 6.4.4 Parasitärer Vertikaler PNP-Transistor

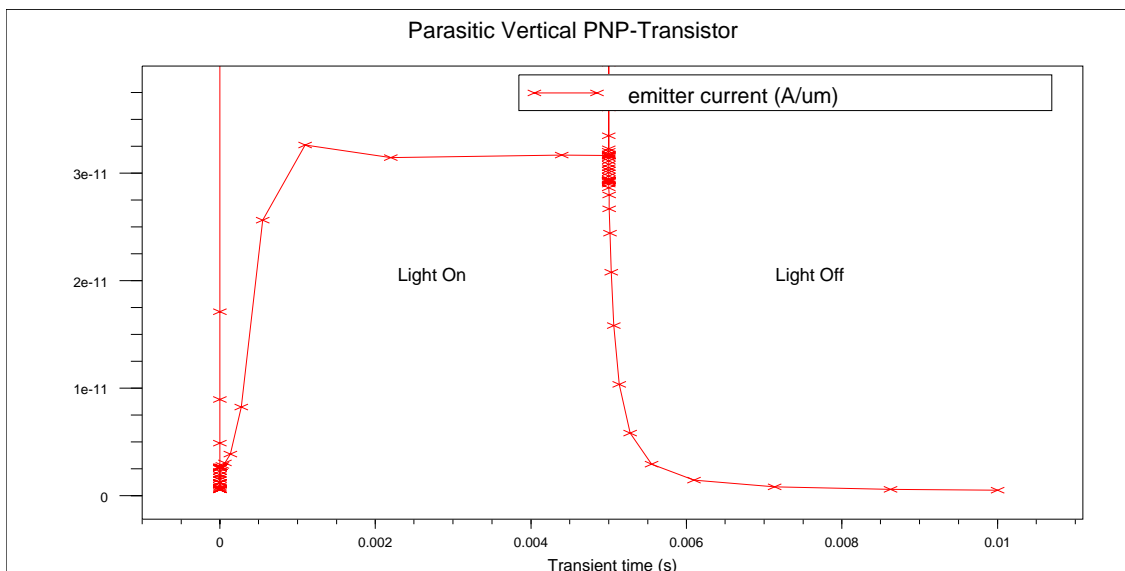


Abbildung 6.28: Emittierstrom in Abhängigkeit der Zeit

Der Parasitäre Vertikale PNP-Transistor hat, wenn er als Phototransistor benutzt wird, als einzigen zugänglichen Anschluß den Emitter, weshalb hier der Emittierstrom aufgetragen ist. Von allen bisher besprochenen Bauteilen weist der PNP-Transistor den höchsten Strom bei Lichteinfall auf. Wie man aber schon an der Zeitachse der Abbildung 6.28 erkennen kann, muß dafür eine längere Schaltzeit in Kauf genommen werden. Der PNP-Transistor erreicht nach ca.  $540\mu s$  den zehnten Teil seines Emittierstromes, der bei Lichteinfall fließt. Die erhöhte Produktion der Ladungsträger, die aufgrund der verstärkenden Wirkung des Transistors erzeugt werden, müssen nach dem Abschalten des Lichtspots wieder rekombinieren. Dies nimmt dann aber eine größere Zeit in Anspruch als dies bei den Dioden der Fall war.

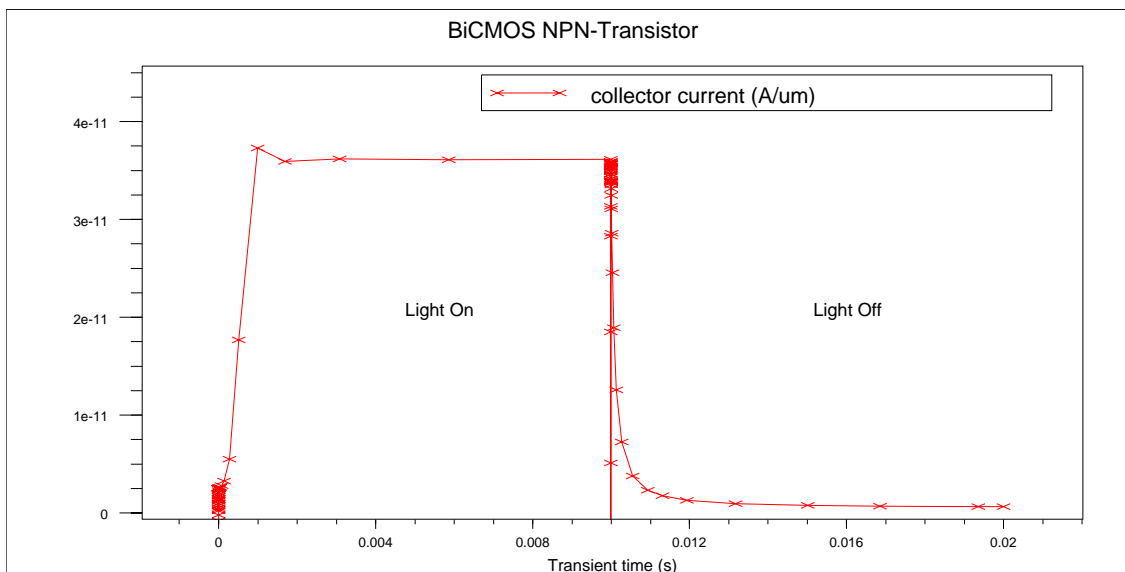


Abbildung 6.29: Kollektorstrom in Abhängigkeit der Zeit

#### 6.4.5 BiCMOS NPN-Transistor

Die Abbildung 6.29 zeigt das zeitliche Verhalten des Kollektorstromes. Hier ist der Kollektorstrom aufgetragen, da der Kollektor einen vom Substrat isolierten Kollektor besitzt. Aufgrund seiner größeren Stromverstärkung  $\beta$  erwartet man einen größeren Strom als beim PNP-Transistor. Das sollte sich jedoch auch in seinen Schaltzeiten bemerkbar machen, die noch höher als beim Parasitären Vertikalen PNP-Transistor sein sollten. Der Kollektorstrom fällt in ca.  $600\mu\text{s}$  auf  $1/10$  des Kollektorstromes bei Lichteinfall ab. Das ist nur wenig langsamer als beim PNP-Transistor. Man muß dabei aber beachten, daß die Stromverstärkung eine Funktion der Basis-Emitter Spannung ist. Bei den hier eingestellten Intensitäten des Lichts kann nur eine geringe Differenz in den Basis- und Emitter-Potentialen erreicht werden. Dort weicht aber die Stromverstärkung des NPN-Transistors erheblich von seinem Maximalwert von ca.100 ab. Der Maximalstrom bei Lichteinfall ist deshalb nur wenig höher als beim PNP-Transistor, ebenso sind die Schaltzeiten nur wenig größer.

Zu den transienten Analysen muß folgendes gesagt werden:

- Beim Zuschalten des Lichtspots steigt der entsprechende Strom in kürzester Zeit sehr stark an, wie das an den Peaks in den Abbildungen 6.25-6.29 zu erkennen ist. Vor allem bei den Dioden kann dieser um einige Größenordnungen höher sein als derjenige, der sich nach einer gewissen Zeit einstellt. Die Zeiten, in denen dies geschieht, sind extrem klein. Diesen extremen Peaks wurden hier keine physikalische Bedeutung (Raumladungskapazitäten o.ä.) zugemessen. Vielmehr wird dies auf numerische Verfahren des Simulators zurückgeführt.
- Analoges gilt beim Abschalten der Spots.
- Die absoluten Werte der Schaltzeiten sind mit großen Unsicherheiten behaftet. Bei

indirekten Halbleitern werden die Lebensdauern der Ladungsträger maßgeblich von sogenannten *Traps*, die in der Mitte der Bandlücke liegen und als Rekombinationszentren dienen, beeinflusst. Die Dichte der Traps und deren genauere energetische Lage in der verbotenen Zone lagen leider bei Abschluß der Arbeit nicht vor.

- Bei allen Bauteilen (ausgenommen der NMOS-Transistor) wurden jedoch die gleichen Modelle und die gleichen Materialeigenschaften verwendet. Die Bauteile unterscheiden sich also nur in der Geometrie bzw. in den Dichteverteilungen der Dotieratome. Sie lassen sich wohl daher *untereinander* relativ gut vergleichen.

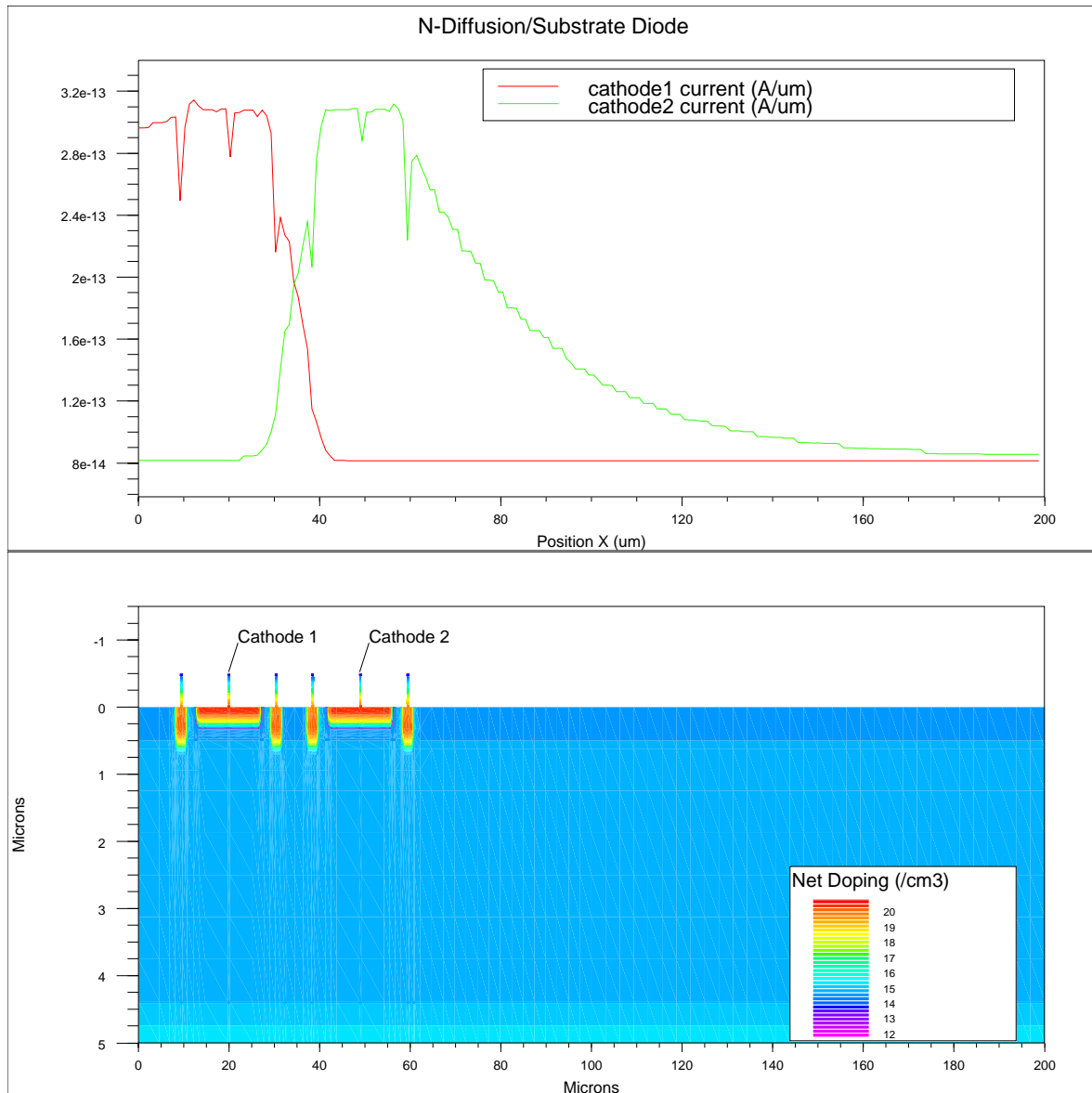
## 6.5 Guard

Im ASIC-Labor Heidelberg wird an verschiedenen Projekten gearbeitet, die zur elektronischen Bilderfassung Pixelarrays, bestehend aus den oben beschriebenen Bauelementen, benötigen. Im *Vision*-Projekt zum Beispiel wurde von Markus Loose eine Pixelkamera, die aus einer Matrix von  $20 \times 20$  N-Diffusion/Substrat Dioden besteht, entwickelt. Idealerweise soll eine Diode nur dann ansprechen (einen Photostrom liefern), wenn ein Lichtstrahl direkt auf sie trifft. Aufgrund seitlicher Diffusion, der vom einfallenden Licht erzeugten Ladungsträger, gelangen diese in den Wirkungsbereich benachbarter Dioden. Es liefern also Dioden, die nicht direkt beleuchtet werden, einen parasitären Strom. Es ist offensichtlich, daß dadurch die Schärfen von Konturen verwaschen werden, das heißt die Auflösung verschlechtert wird.

Weiterhin befindet sich die Ausleseelektronik wie zum Beispiel Ladungsspeicher, Verstärkertextroden etc. ebenfalls auf dem gleichen Chip. Um ohm'sche Verluste der von den Photosensoren gelieferten Ströme möglichst klein zu halten, befindet sich diese dicht neben den optosensitiven Bauelementen. Da sich die „normalen“ Bauelemente physikalisch nicht von den Photosensoren unterscheiden, ist zu erwarten, daß ihre elektrischen Eigenschaften erstens von dem einfallenden Licht und zweitens von den in den Photodioden optisch erzeugten und durch das Substrat diffundierenden Ladungsträger beeinflusst werden. Ersteres könnte durch eine zusätzliche Metallage über der Ausleseelektronik (sofern sie von dem Prozeß zur Verfügung gestellt wird), die das Licht mechanisch abschirmt, abgeschwächt werden. Im zweiten Fall muß man verhindern, daß Ladungsträger in Gebiete diffundieren, wo sie sich durch Verändern der elektrischen Eigenschaften von betroffenen Bauelementen störend bemerkbar machen.

Im Abschnitt 6.2 wurde beim optischen Abscannen der P-Diffusion/Substrat Diode auf das scharfe Abfallen des Anodenstromes hingewiesen. Wenn nämlich der Spot das Gebiet der Wanne verlassen hatte, so wurden die optisch erzeugten Ladungsträger von der Raumladungszone des n-Wannen-Substrat Überganges „abgesaugt“. Laut Simulation gäbe es also zwischen benachbarten P-Diffusion/Substrat Dioden kein Übersprechen. Die N-Diffusion/Substrat Diode erwies sich bezüglich des Übersprechens als anfälliger. Wie man in der Abbildung 6.12 sehen kann fällt der Kathodenstrom relativ schwach ab, wenn der Spot seitlich der Diode auf die Halbleiteroberfläche auftrifft. Bei ihr existiert auch keine zusätzliche pn-Übergangsschicht, die diesen Strom abfangen könnte.

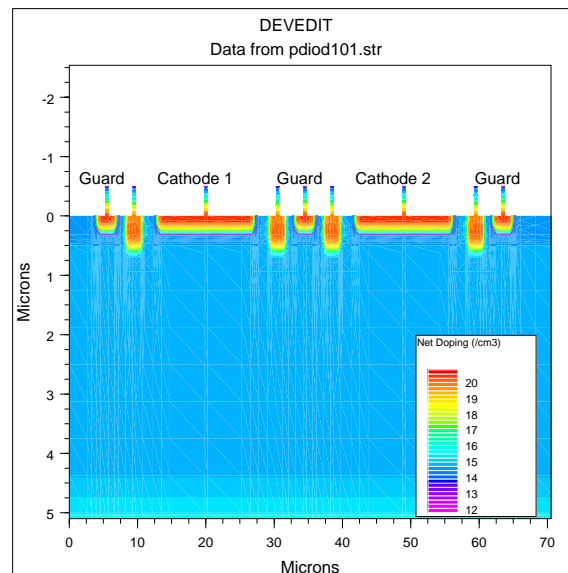
Um das Verhalten zweier benachbarter Dioden zu untersuchen, wurde ein Layout erstellt, das im unteren Teil der Abbildung 6.30 zu sehen ist. Es besteht aus zwei N-Diffusion/Substrat Dioden, die jeweils links und rechts ihrer N-Diffusion zwei Substratkontakte besitzen, die zur Ankontaktierung desselben dienen. Sie wurden zur Übersichtlichkeit im Bild nicht beschriftet.



**Abbildung 6.30:** Kathodenströme zweier benachbarter N-Diffusion/Substrat Dioden beim optischen Abscannen

Damit der Abfall des Kathodenstromes der rechten Diode (Diode 2) beobachtet werden kann, ist der Siliziumblock auf eine Breite von  $200\mu\text{m}$  verbreitert worden. Über diese Strukturen wurde nun ein Lichtspot, der wie in Abschnitt 6.2 definiert ist ( $I = 1\text{W}/\text{m}^2$ ,  $\lambda = 675\text{nm}$ ,  $W = 1\mu\text{m}$ ), mit einer Schrittweite von  $1\mu\text{m}$  bewegt. Die Spannungen betragen bei beiden Kathoden  $+2\text{V}$ , und das Substrat ist auf  $0\text{V}$  gelegt. Der obere Teil des Bildes zeigt die resultierenden Kathodenströme. Wie in Abschnitt 6.2 sieht man deutlich die Einbrüche der Ströme, wenn der Spot über den Aluminiumkontakten liegt. Was aber hier vor allem interessiert, ist der Kathodenstrom der Diode 1, wenn sich der Spot in der Mitte der beiden Dioden bzw. im Gebiet der zweiten Diode befindet. In der Mitte teilen sich die beiden Dioden den Strom, so daß beide gleichermaßen ansprechen. Erst wenn sich der Spot vollständig im Gebiet der zweiten Diode befindet, fällt der Strom der ersten auf den Dunkelwert ab. Betrachtet man den Kathodenstrom 2, so gilt für ihn das gleiche, wenn man von der Diode 2 nach links geht. Nach rechts aber, wo sich keine zusätzliche Raumladungszone befindet, erreicht er seinen Dunkelwert selbst dann noch nicht, wenn der Spot ca.  $140\mu\text{m}$  von der Stelle entfernt ist, wo der Strom der zweiten Diode sein Maximum besaß. Das Licht beeinflußt also die Diode selbst dann noch merklich, wenn es mehr als  $100\mu\text{m}$  entfernt von dieser in das Substrat trifft.

Es scheint also, als ob man zusätzliche Raumladungszonen benötigte, um zu verhindern, daß die Ladungsträger, die weit im Substrat erzeugt werden, zu den Bauteilen diffundieren. In der Abbildung 6.31 wurde daher links und rechts der Dioden und zwischen den beiden Dioden zusätzliche n-Diffusionen eingebaut. Diese werden im Bild als *Guard* bezeichnet. Sie bestehen aus dem gleichen Tiefenprofil, wie die Kathode der Dioden. Die Abstände der Dioden zueinander haben sich nicht verändert. Die Abbildung zeigt nur einen Ausschnitt, in Wirklichkeit ist der Siliziumblock wieder  $200\mu\text{m}$  breit. Die Bezeichnungen der Substratkontakte wurden weggelassen.

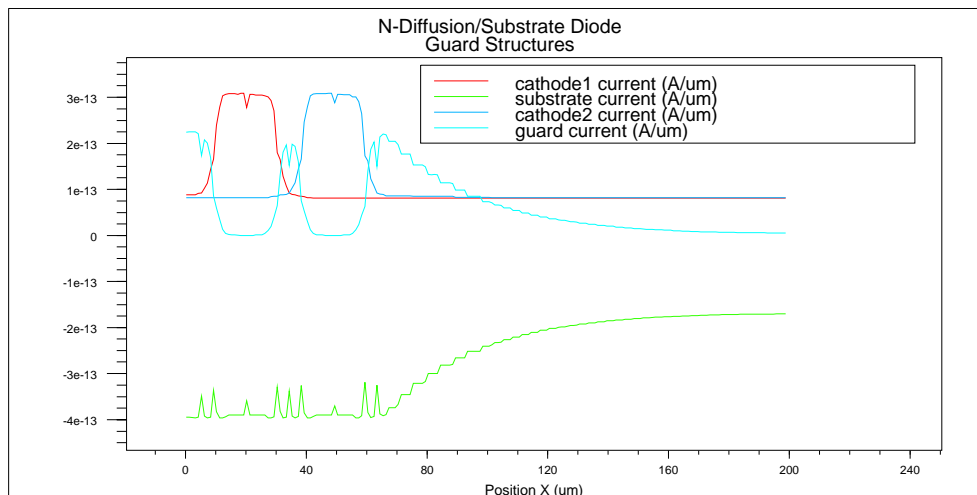


**Abbildung 6.31:** Layout zweier N-Diffusion/Substrat Dioden mit Guard Strukturen

Die drei Guard Strukturen werden zwar kontaktiert, aber sie werden wie das Substrat mit einem externen Potential von  $0\text{V}$  belegt. Die Kathodenspannungen betragen wiederum  $+2\text{V}$ .



Nun wird der gleiche Lichtspot wie oben über den Siliziumblock geführt. Die Abbildung 6.32 zeigt die auftretenden Ströme beim optischen Abscannen. Die rote und dunkelblaue Kurve

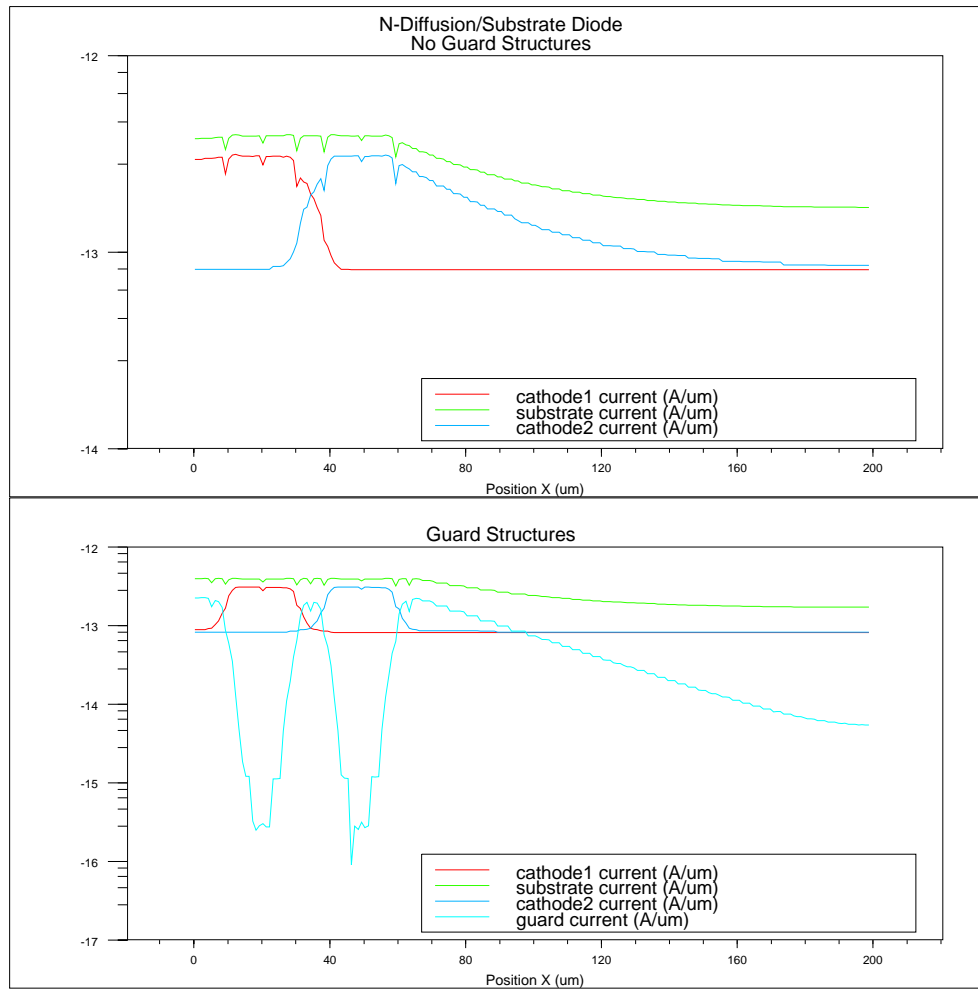


**Abbildung 6.32:** Ströme zweier N-Diffusion/Substrat Dioden mit Guard Strukturen

stellen die Ströme der Kathode 1 und der Kathode 2 dar. Die hellblaue Kurve zeigt den Strom der drei Guard Strukturen, während die grüne Kurve den Substratstrom darstellt. Betrachtet man den Guardstrom, so sieht man, daß er genau das gewünschte Verhalten zeigt: der von den Ladungsträgern erzeugte Strom wird, wenn der Spot neben den Kathoden der Dioden in das Substrat trifft, aufgefangen. An der roten bzw. dunkelblauen Kurve ist deutlich zu erkennen, wie der Kathodenstrom rapide abfällt, wenn das Licht neben den Dioden ins Substrat fällt. Insbesondere auch der Strom der zweiten Diode wird vom Licht, das nur wenige  $\mu\text{m}$  neben der Diode auftrifft, nicht mehr beeinflusst. Laut der Simulation bietet also schon eine relativ kleine n-Diffusion einen guten Schutz vor seitlich in die Diodenregion diffundierenden Ladungsträger, und das, obwohl die Guard Struktur nicht einmal mit einer Sperrspannung belegt wurde.

Zuletzt sollen noch einmal die Kathodenströme der beiden Layouts mit und ohne den Guard Strukturen einander gegenübergestellt werden. Dabei ist zu beachten, daß die Ströme hier auf einer exponentiellen Skala aufgetragen sind. Aufgrund von Diffusion und Lebensdauern der Ladungsträger ergibt sich die Ladungsträgerkonzentration in Abhängigkeit der Entfernung zum Beispiel für die Elektronen zu  $n(x) \propto \exp[-x/L]$ , wobei  $L$  die Diffusionslänge ist, die wiederum proportional der Wurzel der Lebensdauer der Ladungsträger ist. Man kann sehr gut den exponentiellen Abfall des Guardstromes sehen, da diese nicht mit einer Spannung belegt ist. Im Gegensatz dazu fällt der Strom steiler ab, wenn der letzte pn-Übergang (das ist im oberen Bild die Diode2) mit einer Sperrspannung versehen wird.

Was die absolute Größe der Diffusionslängen betrifft, so gilt, da diese durch die Lebensdauern bestimmt wird, das was im vorherigen Abschnitt über die absoluten Größen gesagt wurde.



**Abbildung 6.33:** Vergleich der Ströme zweier N-Diffusion/Substrat Dioden mit und ohne Guard Strukturen

# Kapitel 7

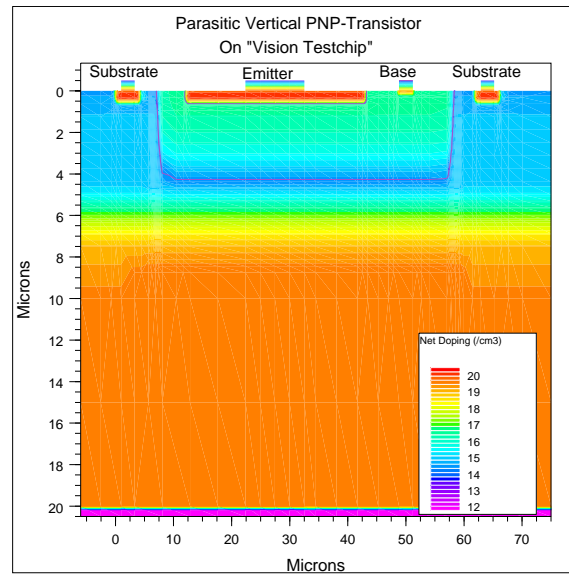
## Messungen

### 7.1 Bipolar

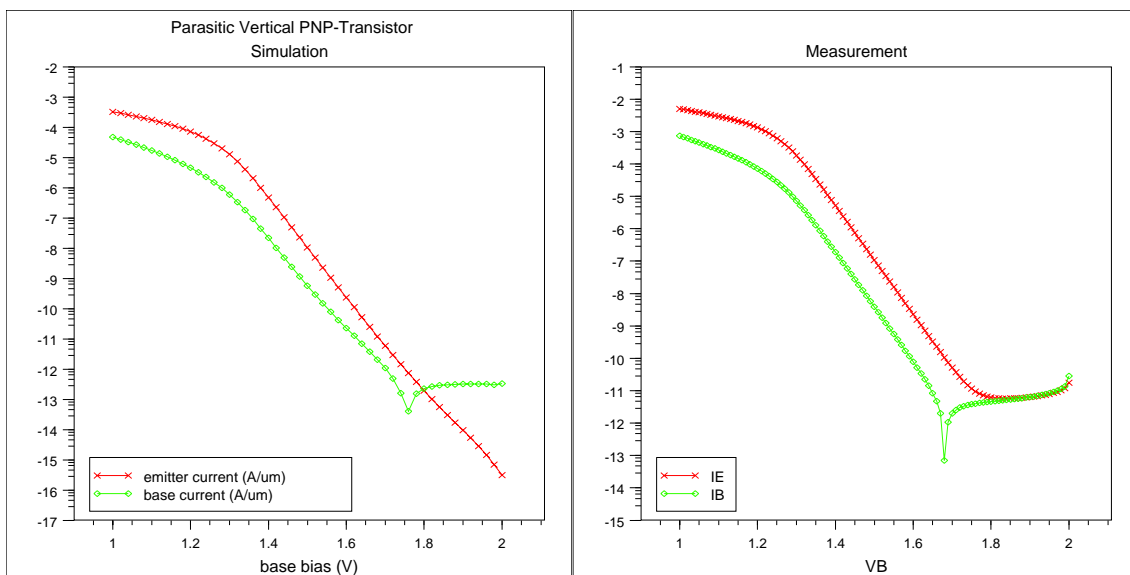
Bei den Simulationen wurden zunächst die vom Hersteller des Simulators gelieferten Standardparameter der entsprechenden Modelle benutzt. Um aber die Bipolar-Simulationen (alle Dioden, der Parasitäre Vertikale PNP- und der BiCMOS NPN-Transistor) den wahren Gegebenheiten anzugleichen, wurden die Lebensdauern der Ladungsträger variiert. Es wurde deshalb dieser Parameter gewählt, da er maßgeblich den Gummel-Plot beeinflusst. Der von "Silvaco International" eingestellte Standardwert beträgt für die Lebensdauern der Ladungsträger  $\tau_{n_0} = \tau_{p_0} = 1 \cdot 10^{-7} s$  (vgl. Kapitel 2). Um eine Übereinstimmung des Gummelplots eines Parasitären Vertikalen PNP-Transistors, der sich auf einem im ASIC-Labor Heidelberg entwickeltem Chip befindet, und dem der Simulation zu erzielen, wurden die Parameter auf  $\tau_{n_0} = \tau_{n_p} = 3 \cdot 10^{-7} s$  eingestellt.

Der auf dem Chip designte Transistor hat nicht die gleichen Maße, wie der in den vorangegangenen Abschnitten benutzte. In der Abbildung 7.1 ist das Layout des Parasitären Vertikalen PNP-Transistors gezeigt, der auf dem Chip realisiert ist, und dessen Design für den Vergleich der Gummelplots benutzt wird. Im Tiefenprofil unterscheidet er sich nicht von dem bisher gezeigten, lediglich der Emitter sitzt nicht mehr in der Mitte der Wanne, und die Basis hat nur einen Kontakt.

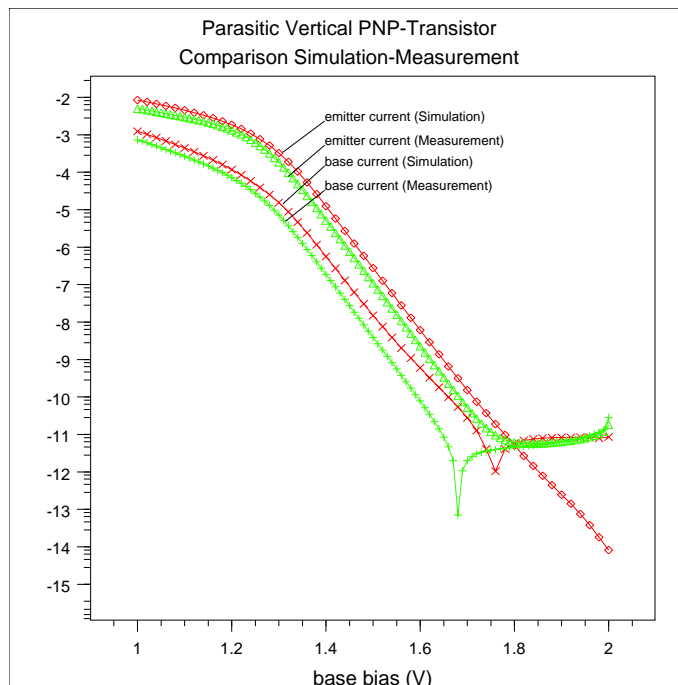
Die Abbildung 7.2 zeigt im linken Teil den Gummelplot der Simulation und im rechten die gemessenen Werte. Hier muß man beim Vergleich beachten, daß der Simulator die Ströme in  $A/\mu m$  angibt. Um also einen echten Vergleich zu haben, müssen die Simulatorwerte mit der, für den Simulator unsichtbaren, Ausdehnung multipliziert werden. Da die Emittergröße maßgeblich in die Verstärkung eingeht, werden die Ströme mit dieser Größe multipliziert. In dieser Richtung hat der Emitter eine Ausdehnung von ca.  $26\mu m$ . Die Abbildung 7.3 zeigt in einer Überlagerung den gemessenen und simulierten Gummelplot. Dabei sind die Stromangaben der Simulation mit  $26\mu m$  multipliziert.



**Abbildung 7.1:** Layout des Parasitären Vertikalen PNP-Transistors nach den Maßen auf dem „Vision Testchip“



**Abbildung 7.2:** Simulierter (links) und gemessener PNP-Transistor

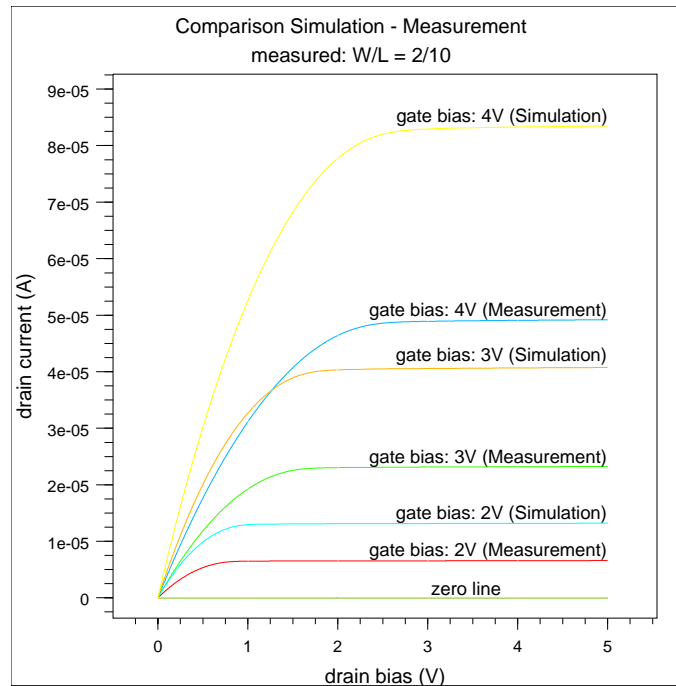


**Abbildung 7.3:** Überlagerung der Gummelplots von simuliertem und gemessenem PNP-Transistor

## 7.2 NMOS

Bei den NMOS-Simulationen wurden ausschließlich die vom Simulator voreingestellten Standardparameter benutzt. Die Abbildung 7.4 zeigt eine Überlagerung der Ausgangskennlinienfelder der Simulation und eines gemessenen NMOS-Transistors. Der gemessene NMOS hat ein  $W/L$ -Verhältnis von  $2/10$ . Die Kanallänge von  $10\mu\text{m}$  entspricht der des im Abschnitt 5 vorgestellten Transistors. Die Weite  $W$  des Transistors muß aufgrund der Zweidimensionalität der simulierten Ströme dazumultipliziert werden. In der Abbildung wurde das berücksichtigt, weshalb der Drainstrom in  $A$  angegeben ist.

Die Kennlinien weichen aber um einen Faktor von  $\approx 2$  voneinander ab. Hier machen sich Randeffekte bemerkbar. Intern rechnet der Simulator mit einem „unendlichen“  $W$ . Sind die Ströme berechnet, so werden sie auf ein  $W$  von  $1\mu\text{m}$  normiert. Daher gehen Randeffekte, bezüglich der Weite  $W$ , nicht in die Simulationsergebnisse ein. Sie machen sich aber bei dem kleinen  $W$  von  $2\mu\text{m}$  des gemessenen Transistors störend bemerkbar. Vergleicht man jedoch die Sättigungsströme mit den Werten, die die Chip-Herstellerfirma *Austria Mikro Systeme International AG* (AMS) für einen NMOS-Transistor mit einem  $W/L$ -Verhältnis von  $100/10$  angibt, so ergibt sich eine Übereinstimmung innerhalb der angegebenen Fehlergrenzen. AMS liefert für alle Bauteile mehrere Parametersätze, die das Verhalten dieser beschreiben. Der *typical mean*-Parametersatz (TM) ist derjenige, der die typischen Werte der produzierten Bauelemente angibt. Da bei der Herstellung der Chips zwischen verschiedenen Proben Differenzen (wie zum Beispiel Abweichungen in den Größen der Belichtungsmasken) auftreten, werden sogenannte „worst case“ Parametersätze von AMS angegeben. Der *worst case speed-*



**Abbildung 7.4:** Überlagerung der Ausgangskennlinienfelder der Simulation und der Messung (Messung von D. Droste)

Parametersatz (WS) gibt den schlechtesten Fall für das Umladen von Kapazitäten an. Dies braucht eine um so längere Zeit, je weniger Strom die Transistoren liefern. Die Parameter im Satz *worst case power* (WP), geben den schlechtesten Fall für den Stromverbrauch an. Die Sättigungsströme sind daher höher, als die des TM-Satzes. Wird die Schaltung mit allen Parametersätzen durchsimuliert, so kann der Designer sicher sein, daß die Funktionalität des Chips, unabhängig von Abweichungen des Prozesses, gewährleistet ist. In der Tabelle 7.1 sind Drainsättigungsströme von drei Parametersätzen und die Ergebnisse der Simulation aufgeführt. Um die Ströme besser miteinander vergleichen zu können, wurden sie alle auf ein  $W$  von  $1\mu m$  normiert. In der letzten Spalte der Tabelle sind die Simulationsergebnisse

Gate Potential [V]	$I_{D_{sat}}$ (WS) [ $10^{-5} A/\mu m$ ]	$I_{D_{sat}}$ (TM) [ $10^{-5} A/\mu m$ ]	$I_{D_{sat}}$ (WP) [ $10^{-5} A/\mu m$ ]	$I_{D_{sat}}$ (Simulation) [ $10^{-5} A/\mu m$ ]
2	$\approx 0.4$	$\approx 0.55$	$\approx 0.74$	$\approx 0.65$
3	$\approx 1.3$	$\approx 1.75$	$\approx 2.24$	$\approx 2.03$
4	$\approx 2.7$	$\approx 3.5$	$\approx 4.39$	$\approx 4.17$

**Tabelle 7.1:** Vergleich der Simulationsergebnisse mit AMS-Parametern

des NMOS-Transistors eingetragen. Man sieht, daß die Daten der Simulation nun besser mit den TM-Werten eines  $W/L = 100/10$ -Transistors übereinstimmen, da die Randeffekte relativ gesehen nicht so stark ins Gewicht fallen. Sie weichen aber immer noch davon ab. Es ist aber auch zu erkennen, daß sie sich innerhalb der angegebenen Grenzen der WS- und

WP-Parameter bewegen. Die Simulation, mit den von der Simulator-Firma voreingestellten Standardparameter, wurde daher als erfolgreich angesehen.





## Zusammenfassung

Der Bauelemente-Simulator hat sich als hilfreiches Werkzeug erwiesen, um einen inneren Einblick in die Funktionsweise von verschiedensten Halbleiterbauelementen zu erhalten. Dabei sind dem "Spieltrieb" nahezu keine Grenzen gesetzt. Das elektrische Verhalten von Bauelementen kann, unter Variation der Geometrien, wie zum Beispiel Emitterbreiten von Transistoren, oder Modifikationen in Höhe und räumlicher Verteilung der Dotierkonzentrationen, mit Hilfe des Simulators untersucht werden. Speziell die hier verwendete *Device Simulation Software V4.0* der Firma "Silvaco International" bietet die Möglichkeit, Halbleitermaterialien wie Silizium und Germanium, ebenso wie zweikomponentige Verbindungshalbleiter, wie Galliumarsenid ( $Ga_xAs_{1-x}$ ), bis zu vierkomponentigen Halbleitern der Form  $In_{1-x}Ga_xAs_yP_{1-y}$ , für das Erstellen von Bauelementen, zu verwenden. Darüber hinaus erlaubt sie den Einsatz verschiedenster Dotiermaterialien, wie Phosphor, Arsen, Antimon und Bor. In dieser Arbeit wurde sich, in Anlehnung an die im ASIC-Labor Heidelberg entwickelten Halbleiterchips, auf Silizium beschränkt. Zusätzlich können mit diesem Programmpaket die Bauelemente mit elektromagnetischen Wellen im sichtbaren Bereich bestrahlt werden, um ihr optisches Verhalten zu untersuchen.

Im Kapitel 3 wurde zunächst ein wenig realistischer pn-Übergang erstellt, um die Ergebnisse der Simulationen mit theoretischen Berechnungen vergleichen zu können. Diese können nur mit erheblichen Vereinfachungen, wie zum Beispiel die einer abrupten Dotierung, durchgeführt werden. Danach wurden die Standarddioden des CMOS-Prozesses untersucht, und ihr unterschiedliches Verhalten den idealisierten Beobachtungen gegenübergestellt. Außer Vergleichen in den Kennlinien der verschiedenen Dioden, wurden die elektrischen Felder, die sich über dem pn-Übergang ausbilden, mit Hilfe von farblich abgestuften Konturplots dargestellt.

Als nächste Erweiterung des pn-Überganges wurden im Kapitel 4 zwei verschiedene Bipolar-Transistoren betrachtet. Dabei wurden die Ladungsträgerverteilungen in den Bauteilen, bei verschiedenen Spannungszuständen, ebenfalls mittels der Konturplots visualisiert. Um die Funktion von Bipolar-Transistoren zu illustrieren, wurden beim BiCMOS NPN-Transistor sowohl Höhe als auch räumliche Ausdehnung der Basiszone modifiziert, wobei zu erkennen war, daß bei ungeeigneter Dotierung die verstärkende Wirkung der Transistoren verloren geht.

In Kapitel 5 wurde versucht die Funktionsweise von MOS-Transistoren zu verstehen. Insbesondere wurde die Ausbildung des Kanals als Funktion des Gate-Potentials dargestellt. Weiterhin ist die Kanallängenmodulation des NMOS auf die zunehmende Ausbreitung der

Verarmungszone, bei steigender Sperrspannung des Drain-Substrat pn-Überganges, in die Kanalregion, zurückgeführt worden.

In Kapitel 6 wurden alle vorher besprochenen Bauteile optischen Einflüssen verschiedenster Art unterworfen. Insbesondere ist dargestellt worden, wie sich Transistoren unter Bestrahlung verhalten, und es wurden die physikalischen Auswirkungen des Lichtes auf diese Bauelemente erklärt. Dabei wurde auch festgestellt, daß der NMOS-Transistor als optischer Sensor ungeeignet ist, da einfallende Photonen lediglich Ströme durch die Raumladungszonen des Drain- und Source-Substrat Überganges erzeugen.

Die Simulationen der Quanteneffizienzen zeigten, daß abhängig von der Tiefe der Raumladungszonen und Höhe der Dotierung, Photodioden Photonen verschiedener Wellenlängen, unterschiedlich gut in Ströme umwandeln können. Dabei sind die Ergebnisse der absoluten Größen der Effizienzen weniger aussagekräftig, als der Vergleich der verschiedenen Dioden untereinander.

Das gilt ebenso für die transienten Simulationen. Die relativen Unterschiede in den Schaltzeiten der verschiedenen Bauteile (Dioden und Transistoren) jedoch, sind durchaus plausibel und decken sich mit anderen Beobachtungen.

Der letzte Abschnitt des Kapitels 6 zeigt eine Möglichkeit auf, Bauteile vor diffundierenden Ladungsträger, die außerhalb deren Bereiche erzeugt werden, zu schützen. Dazu haben sich – in der Simulation – schon relativ kleine und einfache Strukturen als nützlich erwiesen. Um die Resultate zu verifizieren, sind auf einem von D. Droste entwickelten Chip Teststrukturen erzeugt worden. Diese lehnen sich an die Strukturen der Simulation an, so daß Messungen die gemachten Aussagen überprüfen können.

Im letzten Teil der Arbeit schließlich, sind die simulierten Ergebnisse des Parasitären Vertikalen PNP-Transistors und die des NMOS-Transistors mit Messungen und Angaben der Chip Herstellerfirma verglichen worden. Dabei hat sich eine relativ gute Übereinstimmung herausgestellt, so daß die Simulationen durchaus geeignet sind, realistische Prognosen zu stellen.

Es soll hier noch erwähnt werden, daß von den besprochenen Bauteilen Animationen erzeugt wurden, die die Ladungsträgerkonzentrationen in aufeinanderfolgenden äußeren Spannungsschritten zeigen. Im Anhang D sind die dafür benötigten Werkzeuge aufgeführt.

Schließlich sollen noch einige Bemerkungen zum Simulator an sich gemacht werden.

Der hier benutzte Simulator besteht im wesentlichen aus drei Programmteilen:

- “DevEdit“ – eine graphische Oberfläche, mit der die Geometrien und Materialien der Bauteile festgelegt werden
- “Deckbuild“ – ein Editor, in dem die physikalischen und numerischen Modelle angegeben, sowie die Kontakte und äußere Spannungen definiert werden
- “TonyPlot“ – ein Werkzeug zur Darstellung der simulierten Ergebnisse

Im “Device-Editor“ wird, nachdem das Bauteil erstellt ist, ein sogenanntes *mesh* über dieses gelegt. Dieses bestimmt die Stellen, an denen der Simulator die physikalischen Größen berechnet. Die Dichte der Knoten wird vom Benutzer festgelegt, wobei die Anzahl dieser von der Software beschränkt ist ( $\approx 10\,000$ ). Der Benutzer sollte darauf achten, in den für die Funktionsweise wichtigen Regionen (z.B. pn-Übergangsschichten) eine höhere Dichte an

Knoten zu bestimmen, als in weniger bedeutenden (z.B. tief im Substrat). Das bedeutet aber wiederum, daß man sich beim Erstellen des Netzes darüber klar sein muß, *wo* sich die wichtigen Stellen im Halbleiter befinden. Um zum Beispiel einen NMOS-Transistor zu simulieren, dessen aktive Zone (Kanalregion) sich nur wenige *nm* unterhalb der Oberfläche des Halbleiterchips befindet, muß darauf geachtet werden dort eine hohe Dichte ( $\approx 10/nm$  senkrecht zur Oberfläche) an Knotenpunkten zu erhalten, um den Kanal korrekt modellieren zu können. Aufgrund der Beschränkung der Anzahl der Knoten, kann diese Dichte in dem vorgestellten NMOS-Transistor nicht beibehalten werden ( $O(10^6)$  Knoten). Man muß also verschiedene Regionen mit unterschiedlichen Maschenweiten definieren.

Zum „Deckbuild“ sei hier nur ein auf den Anhang C verwiesen, der eine Datei zur Berechnung einer, mit 2V Sperrspannung betriebenen, N-Diffusion/Substrat Diode enthält.



Anhang A

# Maßstabsgetreue Abbildungen

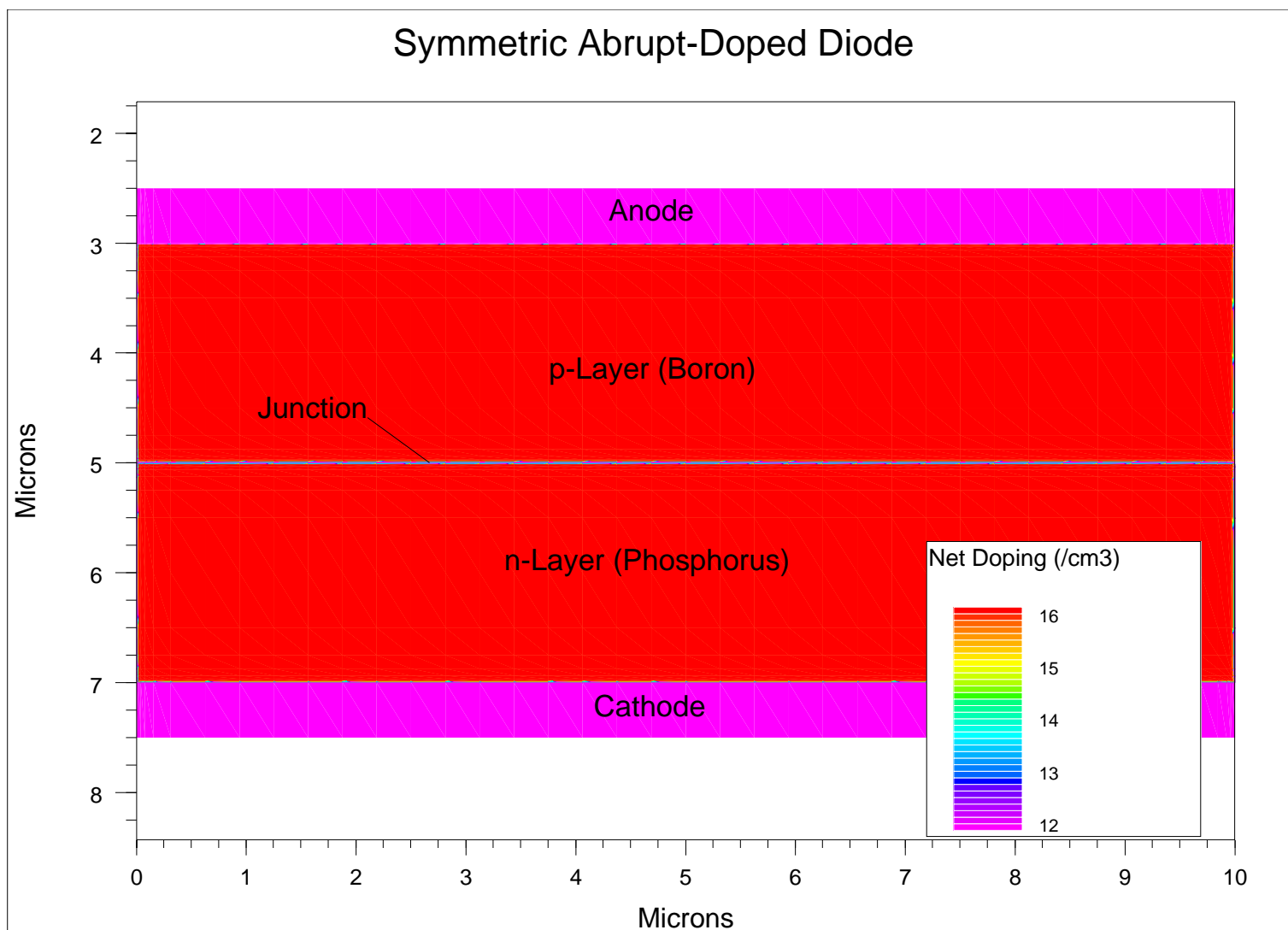
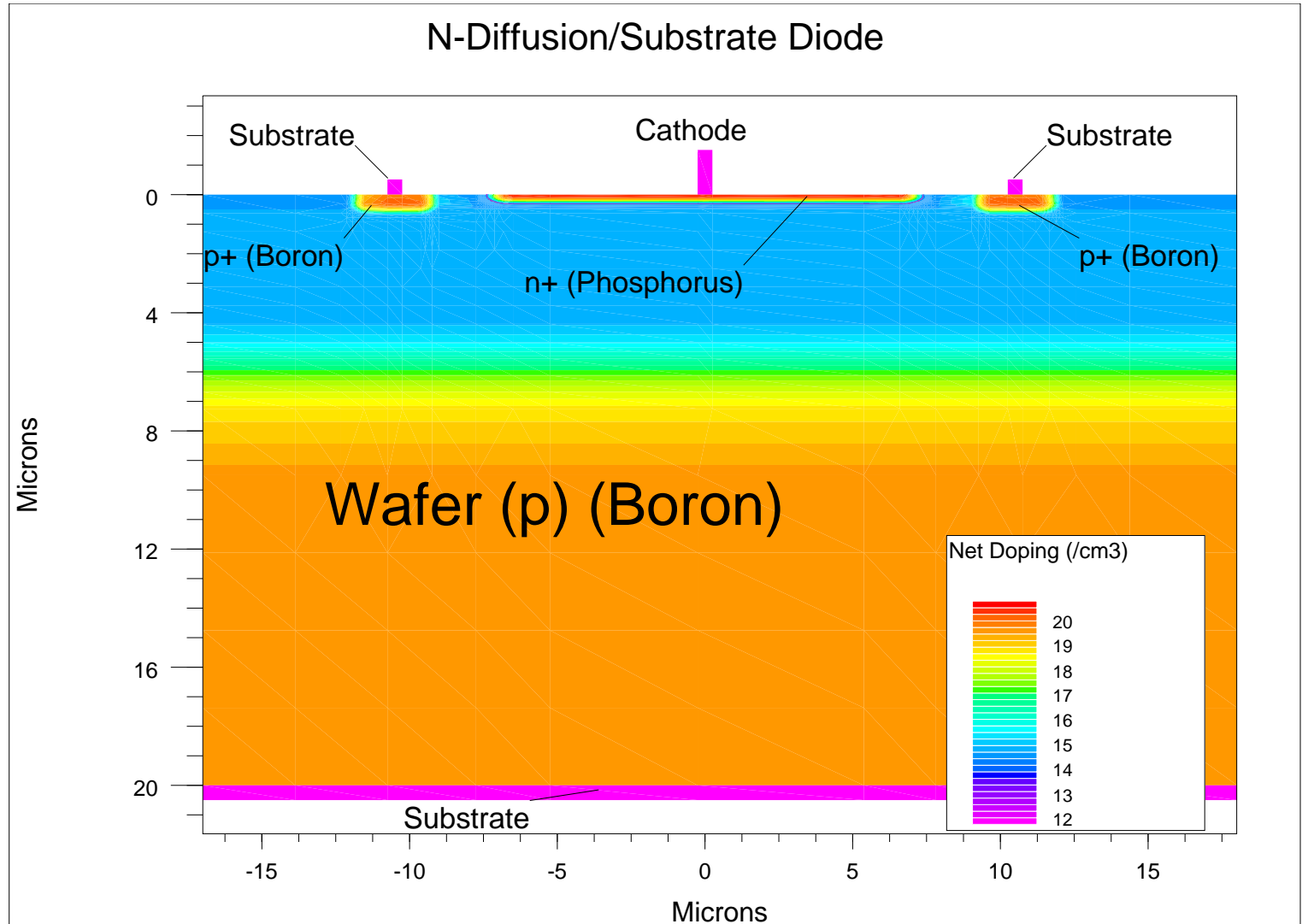


Abbildung A.1: Symmetrisch abrupt dotierte Diode

Abbildung A.2: N-Diffusion/Substrat Diode



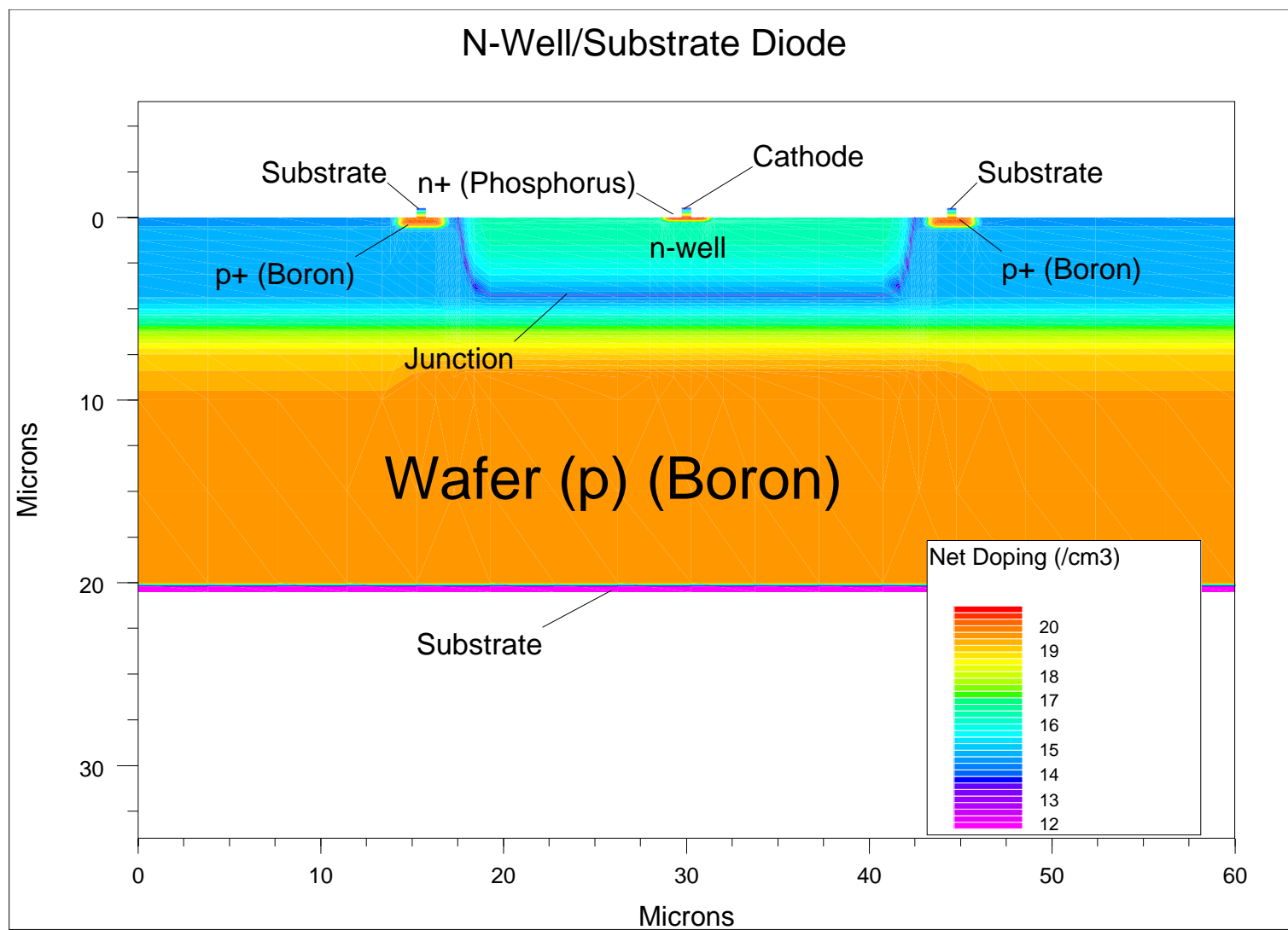
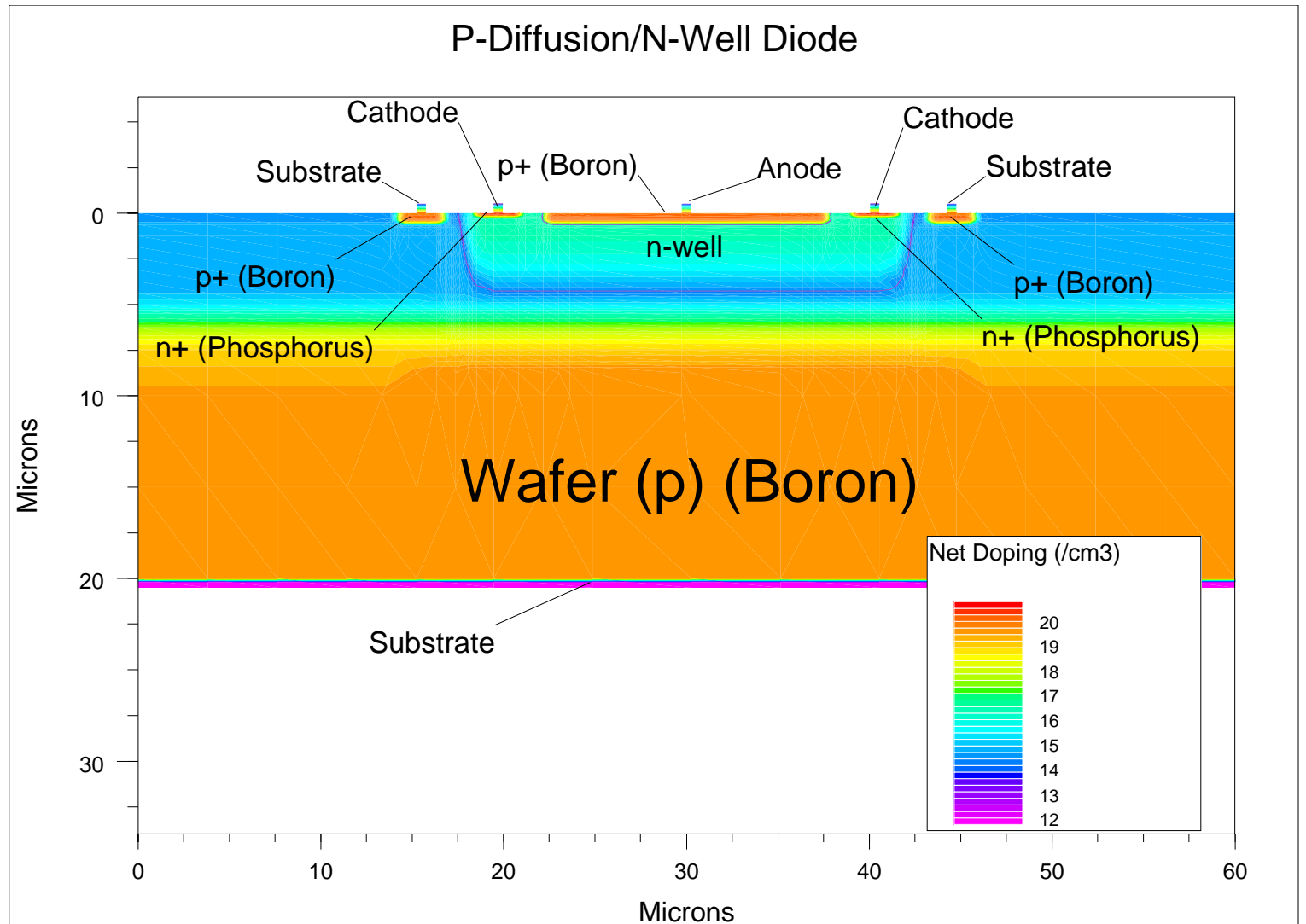


Abbildung A.3: N-Wanne/Substrat Diode



Abbildung A.4: P-Diffusion/N-Wanne Diode



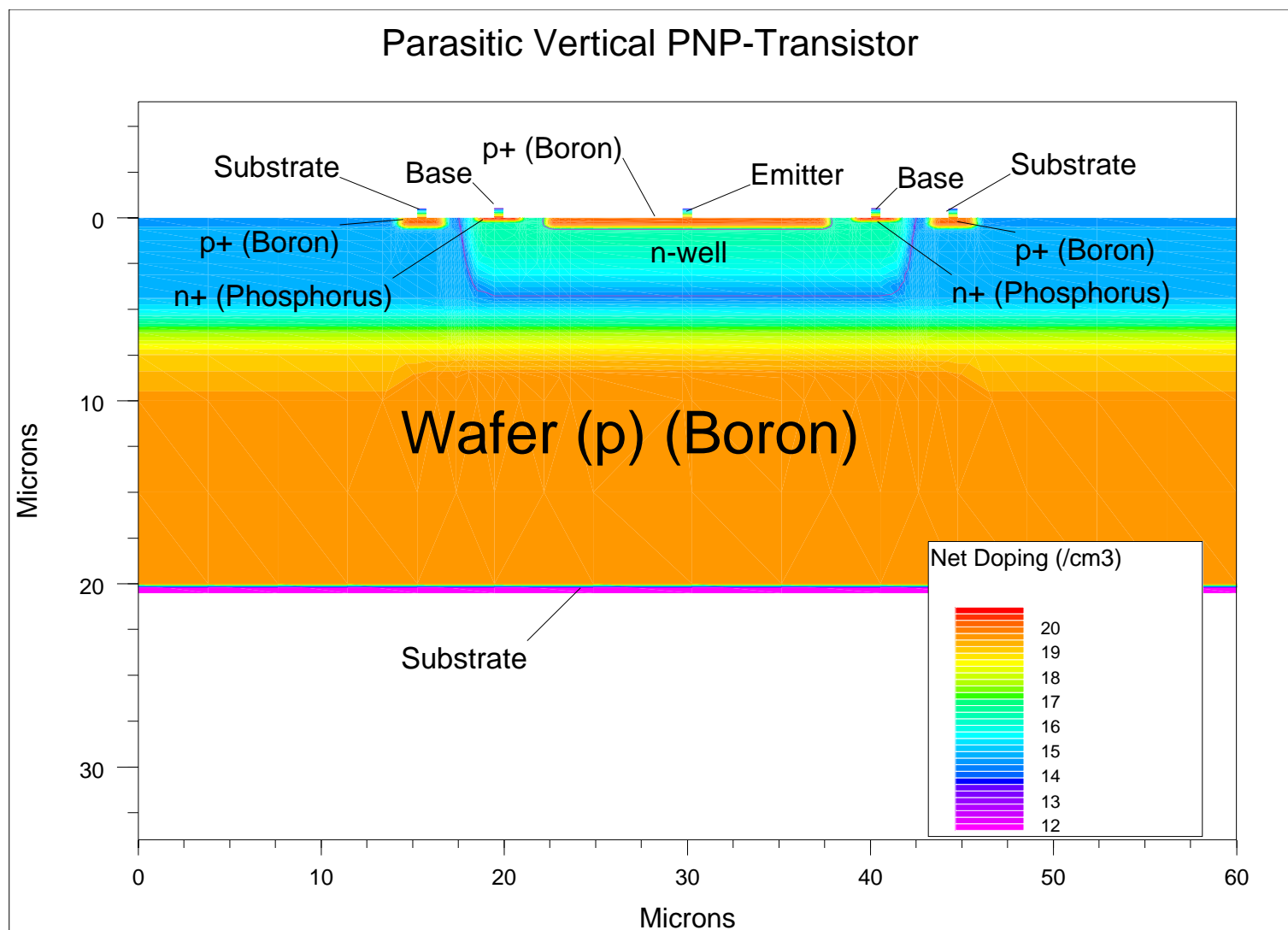
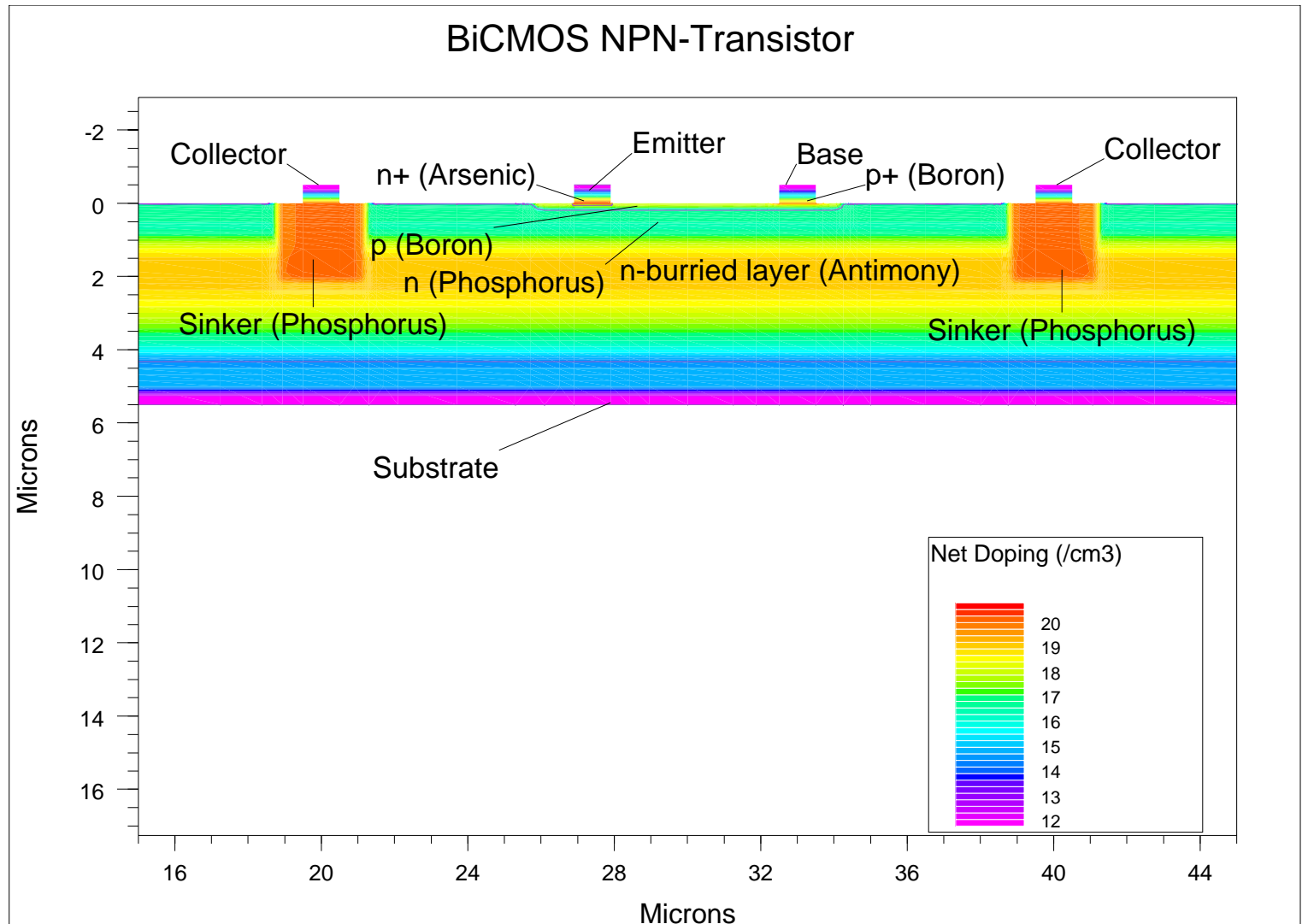
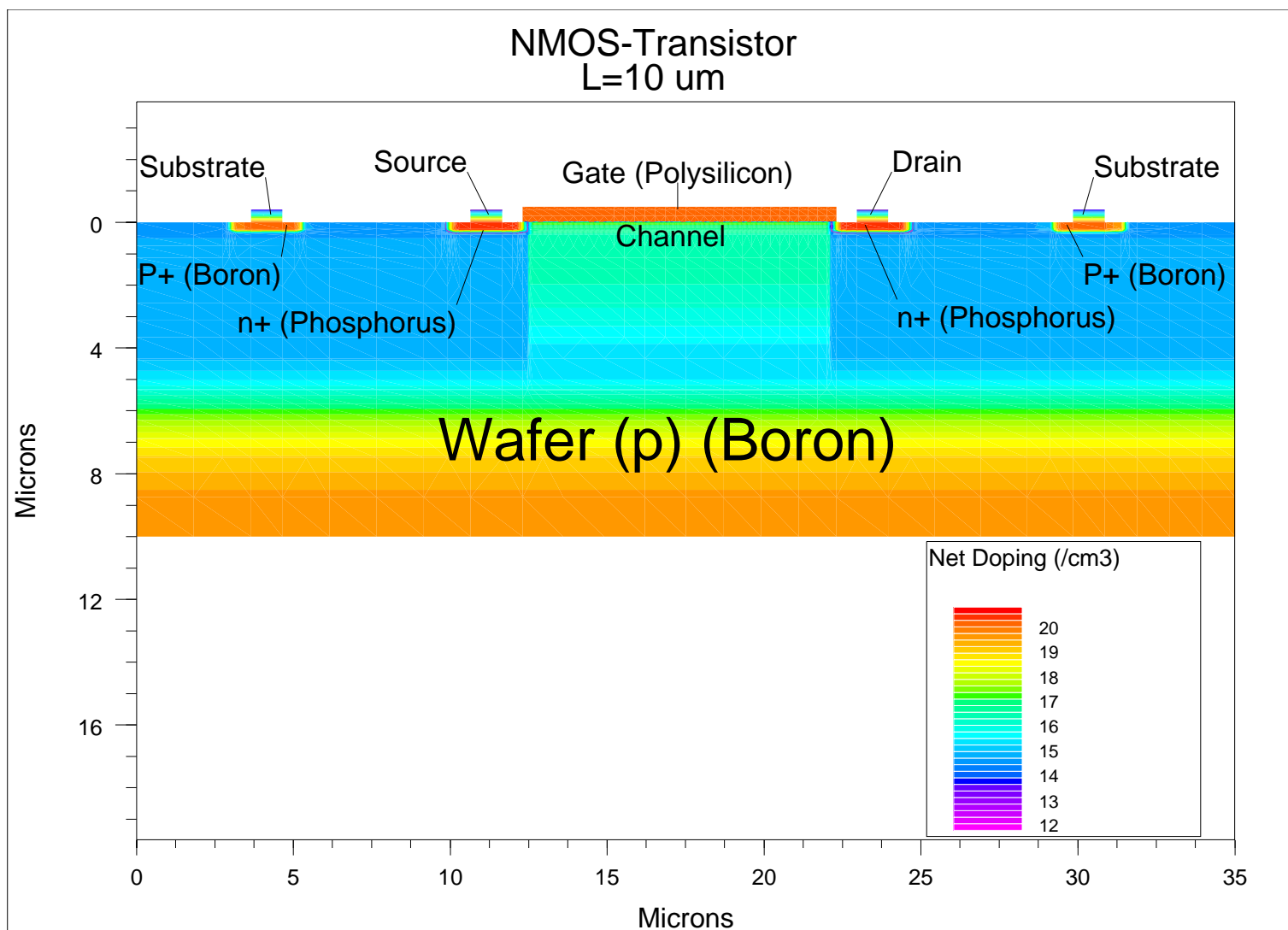


Abbildung A.5: .Parasitärer Vertikaler PNP-Transistor

Abbildung A.6: BiCMOS NPN-Transistor



Abbildung A.7: NMOS-Transistor (Kanallänge  $L = 10\mu\text{m}$ )

# Anhang B

## Elektronenzustände im Festkörper

Bei der Betrachtung von Elektronen im Festkörper ist eines der einfachsten Modelle das des „freien Elektronengases“. Eine Verbesserung besteht darin, das von den Atomrümpfen<sup>1</sup> erzeugte *periodische Potential* zu berücksichtigen. Dieses Modell kann dann erklären, wie es zur Entstehung von *Bändern und Bandlücken (verbotene Zonen)* im Festkörper kommt. Damit läßt sich z.B. auch erklären, warum ein Isolator keine oder keine nennenswerte elektrische Leitfähigkeit besitzt.

### B.1 Freies Elektronengas

Im Modell des freien Elektronengases bewegen sich die Valenzelektronen in einem Kastenpotential, dessen Wände durch die Kristalldimensionen (Kanten  $L$ , Volumen  $L^3$ ) gegeben sind. Ansonsten sind sie aber innerhalb des Kastens *frei* beweglich, das heißt sie wechselwirken weder untereinander („Einelektronennäherung“), noch mit den Atomrümpfen.

Die stationäre Schrödingergleichung für ein „freies“ Elektron im Festkörper-Kastenpotential lautet:

$$-\frac{\hbar^2}{2m}\Delta\Psi(\vec{r}) + V(\vec{r})\Psi(\vec{r}) = E'\Psi(\vec{r}),$$

wobei für ein Kastenpotential gilt:

$$V(x, y, z) = \begin{cases} V_0 = \text{const} & \text{für } 0 \leq x, y, z \leq L \\ \infty & \text{sonst.} \end{cases}$$

Setzt man nun noch  $E = E' - V_0$ , so erhält man für das Innere des Kastens:

$$-\frac{\hbar^2}{2m}\Delta\Psi(\vec{r}) = E\Psi(\vec{r}). \tag{B.1}$$

---

<sup>1</sup>Als *Atomrumpf* wird der von den Valenzelektronen entblößte Rest des Atomes bezeichnet (Kern+Elektronen der inneren Schalen).

Mit den festen Randbedingungen

$$\Psi = 0 \begin{cases} x = 0, L; & 0 \leq y, z \leq L; \\ y = 0, L; & 0 \leq x, z \leq L; \\ z = 0, L; & 0 \leq x, y \leq L; \end{cases}$$

und der Normierungsbedingung  $\int_{V=L^3} \Psi(\vec{r})\Psi^*(\vec{r})d\vec{r} = 1$  ergibt sich die Lösung der Schrödingergleichung zu:

$$\Psi(\vec{r}) = \left(\frac{2}{L}\right)^{3/2} \sin(k_x x) \sin(k_y y) \sin(k_z z).$$

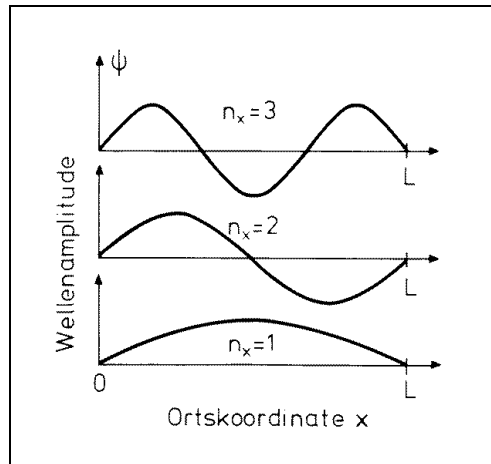
Die möglichen Energieeigenwerte ergeben sich, indem man die Lösung in Gleichung (B.1) einsetzt, zu:

$$E(k) = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2), \quad (\text{B.2})$$

wobei der Wellenzahlvektor  $k$  durch die Randbedingungen folgenden Einschränkungen unterzogen ist:

$$\begin{aligned} k_x &= \frac{\pi}{L} n_x, \\ k_y &= \frac{\pi}{L} n_y, \\ k_z &= \frac{\pi}{L} n_z, \quad \text{mit } n_x, n_y, n_z = 1, 2, 3, \dots \end{aligned}$$

In Abbildung B.1 sieht man die Amplituden der Wellenfunktionen mit den ersten drei Quan-



**Abbildung B.1:** x-Abhängigkeit der ersten drei Wellenfunktionen im rechteckigen Potentialkasten der Länge  $L$  (aus [3])

tenzahlen  $n_x = 1, 2, 3$ . Die zugehörigen Wellenlängen sind  $\lambda_{1,2,3} = 2L, L, \frac{2}{3}L$ .

Wie dicht liegen die Zustände, und wie sind diese mit Elektronen besetzt?

Trägt man die Energieeigenwerte im  $k$ -Raum der Wellenzahlvektoren auf, so erhält man durch die Beziehung  $E = \frac{\hbar^2 k^2}{2m}$  als Flächen gleicher Energie Kugelschalen. Um zur Zustandsdichte zu

<sup>2</sup>sie besagt, daß das Elektron mit Sicherheit *irgendwo* im Kasten anzutreffen ist.

gelangen, muß das Volumen  $V_k = (\pi/L)^3$ , welches ein Zustand im  $k$ -Raum einnimmt, durch das Volumen der Kugelschale, das durch die Flächen  $E(k)$  und  $E(k) + dE^3$  begrenzt wird, dividiert werden<sup>4</sup>. Für die Anzahl der Zustände  $Z$  ergibt sich somit:

$$dZ' = 1/8 \frac{4\pi k^2 dk}{(\pi/L)^3}$$

mit  $dE = (\hbar^2 k/m)dk$  folgt

$$dZ = \frac{(2m)^{3/2}}{4\pi^2 \hbar^3} E^{1/2} dE$$

Es wurde noch nicht berücksichtigt, daß Elektronen einen Spin haben, und somit jeder der Zustände durch zwei Elektronen besetzt werden kann – die Anzahl ist daher doppelt so groß. Die Zustandsdichte  $D(E) = dZ/dE$  beträgt damit<sup>5</sup>:

$$D(E) = \frac{dZ}{dE} = \frac{(2m)^{3/2}}{2\pi^2 \hbar^3} \sqrt{E} \quad (\text{B.3})$$

Diese Zustände werden nun mit Elektronen besetzt. Dabei ist aber zu berücksichtigen, daß die Besetzung der Energiezustände der inneren Energie des Systems entsprechen muß<sup>6</sup>. Es muß also eine temperaturabhängige Verteilungsfunktion  $f(E, T)$  geben, die die Besetzung der Zustände im thermischen Gleichgewicht bestimmt. Die Elektronendichte ergibt sich somit zu:

$$n = \int_0^\infty D(E) f(E, T) dE. \quad (\text{B.4})$$

Da Elektronen Fermionen sind (Spin 1/2-Teilchen) gehorchen sie der *Fermi-Dirac*-Statistik. Die fermi'sche Verteilungsfunktion lautet:

$$f(E, T) = \frac{1}{e^{\frac{(E-\mu)}{kT}} + 1}, \quad (\text{B.5})$$

wobei  $T$  die absolute Temperatur in Kelvin,  $E$  die Energie des betrachteten Zustandes,  $\mu$  das chemische Potential und  $k$  die Boltzmann Konstante bezeichnen. Um die Bedeutung des chemischen Potentials zu sehen, betrachtet man die Verteilungsfunktion für  $T \rightarrow 0$ :

$$f(E < \mu, T \rightarrow 0) \rightarrow 1$$

$$f(E > \mu, T \rightarrow 0) \rightarrow 0$$

$$f(E = \mu, T \rightarrow 0) = 1/2.$$

---

<sup>3</sup>Die Wellenzahlvektoren (und damit die Energien) sind zwar prinzipiell diskret, liegen aber derart dicht, daß von der Summation auf die Integration übergegangen werden kann.

<sup>4</sup>Durch die gegebenen *festen* Randbedingungen liegen die Zustände nur jeweils auf den positiven Halbachsen im  $k$ -Raum; das Volumen beträgt somit nur 1/8 des Kugelvolumens.

<sup>5</sup>Die Zustandsdichte wurde mit *festen* Randbedingungen abgeleitet. Das gleiche erhält man jedoch, wenn man zu *periodischen* Randbedingungen  $\Psi(x+L, y+L, z+L) = \Psi(x, y, z)$  übergeht (siehe z.B. [3]).

<sup>6</sup>Am absoluten Nullpunkt ( $T = 0$ ) werden sie sukzessive von unten nach oben mit Elektronen besetzt.

Die Verteilungsfunktion ist für  $T \rightarrow 0$  also eine Stufenfunktion, die an der Stelle  $E = \mu$  von 1 auf 0 springt. Am absoluten Nullpunkt sind wie oben erwähnt die *untersten* Zustände ( $E < \mu$ ) bis zu einer gewissen Grenze ( $E = \mu$ ) mit Elektronen besetzt. Diese Grenze entspricht aber gerade der *Fermienergie*  $E_F$  bzw. dem *Fermi-niveau*. Das heißt  $\mu(T = 0) = E_F$ .

(frei nach [3])

Die Fermienergie am absoluten Nullpunkt läßt sich mit den nun bekannten Tatsachen aus Gleichung (B.4) berechnen:

$$n = \int_0^\infty D(E)f(E,T)dE = \int_0^{E_F} \frac{(2m)^{3/2}}{2\pi^2\hbar^3} E^{1/2} \cdot 1 dE = \frac{(2m)^{3/2}}{2\pi^2\hbar^3} \frac{2E_F^{3/2}}{3}$$

Weil die Zustände mit  $E < E_F$  *alle* besetzt sind und darüber hinaus keiner mehr, erfolgt die Integration anstatt von 0 bis  $\infty$  von 0 bis  $E_F$ . Außerdem ist in diesem Energiebereich die Verteilungsfunktion  $f(E,T) = 1$ . Löst man nun nach  $E_F$  auf, so erhält man:

$$E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (\text{B.6})$$

In der Tabelle B.1 sind einige Metalle mit ihren zugehörigen Elektronendichten, Fermienergien und Fermitemperaturen aufgeführt. Bemerkenswert ist die Höhe der Fermitemperatur, die Konsequenzen bezüglich der Besetzungszahlen bei endlichen Temperaturen hat.

Metall	$n[10^{22} \text{cm}^{-3}]$	$E_F[\text{eV}]$	$T_F[\text{K}]$
Li	4.62	4.70	54 500
Na	2.53	3.14	36 400
Al	18.07	11.65	135 200
Cu	8.47	7.03	81 600

**Tabelle B.1:** Fermienergien und dazugehörige Fermitemperaturen einiger Metalle (aus [3])

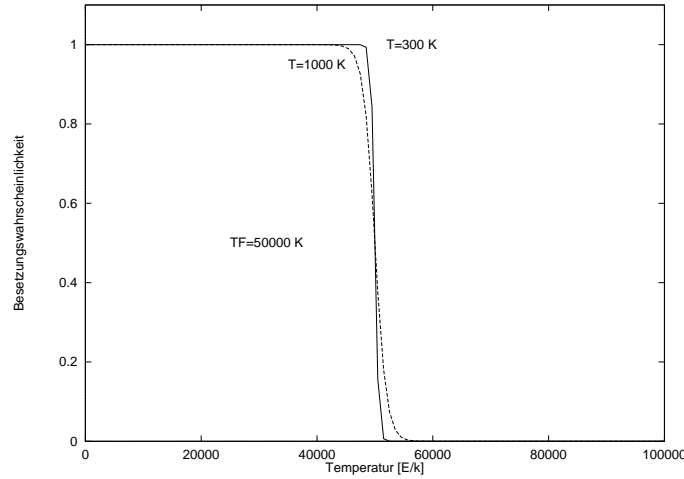
Was passiert nun wenn man zu einer endlichen Temperatur übergeht?

Um zu verdeutlichen, wie die Fermienergie die Besetzung der Zustände bei endlichen Temperaturen beeinflusst, führt man die *Fermitemperatur*  $T_F$  ein:  $T_F = E_F/k$  ( $k = 86 \cdot 10^{-6} \frac{\text{eV}}{\text{K}}$ ). In der Abbildung B.2 sieht man die fermi'sche Verteilungsfunktion für verschiedene Temperaturen. Die Fermi-Temperatur  $T_F$  wurde hier zu  $50\,000\text{K}$  gewählt. Man sieht, daß die Stufe der Fermi-Funktion bei endlichen Temperaturen keine exakte Stufe mehr ist, sondern etwas „aufweicht“. Andererseits sieht man aber auch, daß aufgrund der hohen Fermitemperatur  $T_F$  der Metalle die Aufweichung relativ schwach ist. Das bedeutet, daß sich das freie Elektrogas im Festkörper bei Zimmertemperatur nahezu so verhält wie am absoluten Nullpunkt der Temperatur. Es sind also auch bei endlichen Temperaturen nur sehr wenige Zustände oberhalb der Fermienergie besetzt.

## B.2 Periodisches Potential

Wie kommt es nun zur Bandstruktur in einem Festkörper?





**Abbildung B.2:** Fermi-Verteilungsfunktion bei verschiedenen Temperaturen

Um zu verstehen, wie es zur Ausbildung von Bändern im Festkörper kommt, betrachtet man ein Elektron, das sich unter dem Einfluß eines periodischen Potentials, welches von den Gitteratomen gebildet wird, befindet. Die stationäre Schrödingergleichung lautet dann:

$$-\frac{\hbar^2}{2m} \Delta \Psi(\vec{r}) + V(\vec{r}) \Psi(\vec{r}) = E \Psi(\vec{r}), \quad (\text{B.7})$$

mit dem periodischen Potential  $V(\vec{r}) = V(\vec{r} + \vec{R})$ , wobei  $\vec{R}$  ein Translationsvektor im dreidimensionalen Festkörpergitter bezeichnet. Wegen der Periodizität des Potentials, läßt sich dieses in eine Fourier-Reihe entwickeln:

$$V(\vec{r}) = \sum_{\vec{G}} V_{\vec{G}} e^{i\vec{G}\vec{r}}.$$

Hier ist  $\vec{G}$  ein *reziproker* Gittervektor<sup>7</sup>. Durch die Entwicklung nach reziproken Gittervektoren wird die Periodizität des Potentials durch das Kristallgitter berücksichtigt. Die Wellenfunktion  $\Psi(\vec{r})$  wird nach ebenen Wellen entwickelt:

$$\Psi(\vec{r}) = \sum_{\vec{k}} C_{\vec{k}} e^{i\vec{k}\vec{r}},$$

$\vec{k}$  bezeichnet einen Punkt des reziproken Raumes. Setzt man nun  $\Psi(\vec{r})$  und das Potential  $V(\vec{r})$  in Gleichung (B.7) ein, so folgt:

$$\sum_{\vec{k}} e^{i\vec{k}\vec{r}} \left[ \left( \frac{\hbar^2 k^2}{2m} - E \right) C_{\vec{k}} + \sum_{\vec{G}} V_{\vec{G}} C_{\vec{k}-\vec{G}} \right] = 0.$$

Weil diese Bedingung für jedes  $\vec{r}$  gilt, muß der Ausdruck in der eckigen Klammer für jedes  $\vec{k}$  verschwinden. Wegen der periodischen Randbedingungen ergeben sich bei  $N$  Elementarzellen jeweils  $N$  erlaubte Wellenzahlvektoren als Lösungen in *jeder* Brioullin-Zone. Jede dieser

<sup>7</sup>Im eindimensionalen gilt:  $G = 2\pi h/a$  ( $a$ : Basisvektor der Gitterzelle).

Lösung (festes  $\vec{k}$ ) besteht aus einer Superposition von ebenen Wellen, deren Wellenzahlvektoren sich nur um einen reziproken Gittervektor  $\vec{G}$  unterscheiden. Jedem  $\vec{k}$  entspricht damit auch ein Energieeigenwert  $E(\vec{k})$ .

### B.3 Effektive Masse

Um die Bewegung von Elektronen innerhalb eines Bandes von Festkörpern mit der klassischen newton'schen Bewegungsgleichung  $\vec{F} = \dot{\vec{p}}$  beschreiben zu können, führt man das Konzept der *effektiven Masse* ein. Dazu betrachtet man die Gruppengeschwindigkeit einer Bloch-Welle:

$$\dot{\vec{r}} = \vec{v}(\vec{k}) = \frac{1}{\hbar} \nabla_{\vec{k}} E(\vec{k}) = \frac{1}{\hbar} \frac{\partial E(\vec{k})}{\partial \vec{k}}$$

Dies liefert gerade für das freie Elektron mit der Energie  $E = \frac{\hbar^2 \vec{k}^2}{2m}$  den Zusammenhang zwischen Geschwindigkeit und Impuls  $\vec{v} = \vec{p}/m$ .

Wirkt nun eine äußere Kraft  $\vec{F}$ , so lautet die Bewegungsgleichung:

$$\hbar \dot{\vec{k}} = \vec{F} = -e \left[ \vec{E} + \vec{v}(\vec{k}) \times \vec{B} \right]$$

Aus diesen beiden Gleichungen folgt nun:

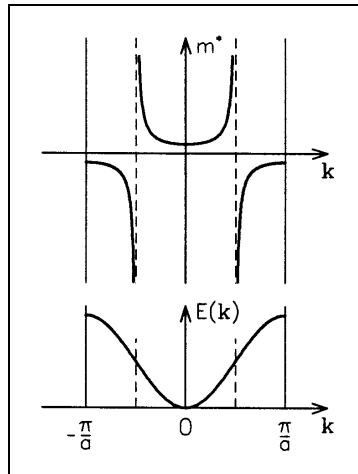
$$\dot{\vec{v}} = \frac{\partial}{\partial t} \left[ \frac{1}{\hbar} \frac{\partial E(\vec{k})}{\partial \vec{k}} \right] = \frac{1}{\hbar} \frac{\partial}{\partial \vec{k}} \left[ \frac{1}{\hbar} \frac{\partial E(\vec{k})}{\partial \vec{k}} \right] \hbar \dot{\vec{k}} = \frac{1}{\hbar^2} \frac{\partial^2 E(\vec{k})}{\partial \vec{k} \partial \vec{k}} \vec{F}.$$

Für die einzelnen Komponenten des Beschleunigungsvektors gilt:

$$\dot{v}_i = \frac{1}{\hbar^2} \sum_j \frac{\partial^2 E(\vec{k})}{\partial k_i \partial k_j} F_j.$$

Um einen analogen Ausdruck für die Bewegungsgleichung zu bekommen setzt man  $\frac{1}{\hbar^2} \frac{\partial^2 E(\vec{k})}{\partial k_i \partial k_j} = \left( \frac{1}{m^*} \right)_{ij}$  und bezeichnet  $m^*$  als *effektive Masse*. Sie ist also durch die zweite Ableitung der Energie  $E(\vec{k})$  nach dem Wellenzahlvektor  $\vec{k}$  bestimmt. Das ist gerade die *Krümmung* der  $E(\vec{k})$ -Kurve. Auch für das Valenzband kann eine effektive Masse eingeführt werden. Damit man für die Valenzbandoberkante ebenfalls positive (Krümmung ist entgegengesetzt) effektive Massen bekommt, setzt man für dieses Band:  $-\frac{1}{\hbar^2} \frac{\partial^2 E(\vec{k})}{\partial k_i \partial k_j} = \left( \frac{1}{m^*} \right)_{ij}$ . Im Gegensatz zur „normalen“, newton'schen Masse ist die effektive Masse im allgemeinen keine skalare Größe<sup>8</sup>, sondern weist Richtungsabhängigkeit auf, das heißt sie ist ein Tensor. Dieser ist aber symmetrisch und läßt sich auf die Hauptachsen transformieren, so daß man im höchsten Fall drei Komponenten hat. Anhand der Definition der effektiven Masse läßt sich erkennen, daß Elektronen in Zuständen, in denen die Krümmung der Energiefläche groß ist, die Masse klein ist und umgekehrt. Sie geht an den Stellen gegen unendlich, an denen die Krümmung der Energieflächen null wird. Abbildung B.3 zeigt eine  $E(\vec{k})$ -Kurve und die dazugehörige effektive Masse  $m^*$ . Ebenso ist in der Abbildung zu erkennen, daß in der Nähe der Extrema die effektive Masse nahezu konstant ist.

<sup>8</sup>Bei Brioullin Zonen mit Kugelgestalt ist sie ein Skalar.



**Abbildung B.3:** Effektive Masse und Dispersionsrelation (aus [7])

In den Leitungsbandmaxima ist die effektive Masse sogar negativ, was bedeutet, daß sich die Ausbreitungsrichtung der Elektronen dort umkehrt (*Bragg-Reflexion*). Gäbe es also sonst keine Störungen (Verunreinigungen, Phononen, Stöße mit anderen Elektronen, ...) der Bewegung der Elektronen, dann dürfte es in makroskopischen Kristallen keine Gleichstromleitfähigkeit geben, da die Elektronen an der Grenze der Brillouin-Zone reflektiert werden. (nach [7])

Betrachtet man die Bewegung der Elektronen im *fast* voll besetzten Valenzband, dann müßte man deren effektive Masse über alle Krümmungen des Energiebandes mitteln. Es ist daher einfacher, die Bewegung des *Defektelektrons*, oder auch *Loches*, zu untersuchen. Wenn ein Elektron vom Valenzband ins Leitungsband wechselt, so bleibt im Valenzband ein Loch mit positiver Ladung zurück. Da vornehmlich Elektronen aus dem *Maximum* der Valenzbandkante ins Leitungsband wechseln und dort die effektive Masse nahezu konstant ist, kann wieder mit einer einheitlichen effektiven Masse, jetzt aber für das verbleibende Loch im Valenzband, gerechnet werden. Hier haben also die Elektronen im Leitungsband und die Löcher im Valenzband effektive Massen *gleichen* Vorzeichens. Aufgrund der entgegengesetzten Richtungen der Ausbreitungsgeschwindigkeit und der entgegengesetzten Ladung (Elektron:  $-e_0$ , Loch:  $+e_0$ ) tragen beide gleichermaßen zum Stromtransport bei.

# Anhang C

## Deckbuild-Datei

```
# Photodiode
#-----
# ndiff.src
#-----
go atlas
mesh infile=pdiod112.str master.in
# - models and settings
material material=Silicon taun0=3e-7 taup0=3e-7
models auger conmob consrh fldmob b.electrons=2 b.holes=1 evsatmod=0
hvsatmod=0 boltzman bgn print numcarr=2 temperature=300
contact num=1 name=substrate
contact num=2 name=cathode
contact num=3 name=substrate
contact num=4 name=substrate
method newton autotr trap
#
# -initialisation
solve init
# -make xy-graph
log outfile=ndiffbeam.log
# -ramp cathode
solve vcathode=0 vstep=0.1 vfinal=2 name=cathode
# -optical source (Lightspot)
beam num=1 wavelength=0.675 x.o=0 y.o=-2 ang=90 min.w=-5 max.w=5
solve b1=1e-7
solve b1=1e-6
solve b1=1e-5
solve b1=1e-4
# -save the final structure
save outfile=ndiffbeam.str
quit
```

Zeilen, die mit # beginnen sind Kommentarzeilen und werden vom Simulator ignoriert. Zunächst wird ein mit DevEdit erstelltes Bauelement geladen und die Materialeigenschaften, die physikalischen Modelle, die Anschlußkontakte und die numerischen Methoden definiert. Dann werden alle Kontakte zunächst auf 0V gebracht (solve init) und anschließend die Kathode auf 2V hochgefahren, wobei in dem ".log" -Datei die Ströme in Abhängigkeit der Kathodenspannung aufgezeichnet werden. In den nächsten Zeilen wird der Lichtstrahl definiert, und mit einer Intensität von  $10^{-4}W/cm^2$  auf die, mit 2V Sperrspannung betriebene Diode gestrahlt. Zum Schluß wird in Konturplot-Format das Ergebnis abgespeichert.

# Anhang D

## Skript zur Animation

```
echo
echo ++++++[PS 09/1997]+
echo +++.....MOVIE.....+++
echo ++++++
#(Hilfe: /cad4/caduser/pschneid/silvaco/movie/read.me)
rm $1.ps
rm $1anim.gif
#-----
#— analysiert deckbuild input file und erzeugt outputfiles
#— input: deckbuildfile fn.src
# setfile fn.set
#— output: fn.srm
# fn.set
# fn.dbc
# fn.bsh
#
#echo deckbanalyze.exe $1 ...
/bin/deckbanalyze.exe $1
#
echo vsolve vstep vfinal Nr filename
#
while test -r $1.bsh
do
#
*****
#
cat $1.dbc
#
#-----
#— konvertiert einzelne deckbuildfiles und setfiles
# entfernt fn.bsh wenn vsolve=vfinal
#... input: fn.srm
# fn.set
# fn.dbc
#... output: fn.srm
# fn.set
# fn.dbc
#
#echo deckbconv.exe $1 ...
/bin/deckbconv.exe $1
#
#-----
#— startet simulator atlas
#... input: fn.srm
# fn.set
#... output: fn.ps (= structure file fuer tonyplot)
#
#echo deckbuild -run $1.srm -outfile $1.out ...
deckbuild -run $1.srm -outfile $1.out
#
#-----
#— konvertiert postscript in gif
#... input: fn.ps
#... output: fn.gif
#
#echo pstogif $1.ps ...
pstogif $1.ps
#
#-----
#— erzeugt gif-filenames mit laufender nummer (max.: 9999)
#... input: fn.gif
#... output: fnxxx.gif
```

```
#
#echo lfdnrconv.exe $1 ...
/bin/lfdnrconv.exe $1
#
#-----
#— entfernt fn.ps
#... input: fn.ps
#... output: -/-
#
#echo rm $1.ps ...
rm $1.ps
#
#*****
#
done
#
#-----
#— erstellt aus .gif-files animated gif
#
gifmerge $1*.gif -10 > $1anim.gif
```





# Literaturverzeichnis

- [1] T. Mayer-Kuckuck. *Atomphysik*, Teubner Studienbücher, Stuttgart, 4. Auflage 1994.
- [2] K. Kopitzki. *Einführung in die Festkörperphysik*, Teubner Studienbücher, Stuttgart, 2. Auflage 1989.
- [3] Ibach, Lüth. *Festkörperphysik*, Springer Lehrbuch, Berlin, 4. Auflage 1995.
- [4] Bergmann, Schäfer. *Bd.6 Festkörperphysik*, de Gruyter, Berlin, 1992.
- [5] Ch. Kittel. *Einführung in die Festkörperphysik*, R. Oldenbourg, München, 3. Auflage 1973.
- [6] S. M. Sze. *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1985.
- [7] S. Hunklinger. *Festkörperphysik Vorlesungsskript*, Universität Heidelberg, WS 94/95.
- [8] Silvaco International. *ATLAS User's Manual Device Simulation Software*, Santa Clara USA, Version 4.0 1995.
- [9] H. R. Lacour. *Elektronische Bauelemente Bd. 2*, Berliner Union, Stuttgart 2. Auflage 1980
- [10] IEEE Transactions on Electron Devices. *Measurement of Collector and Emitter Resistances in Bipolar Transistors*, vol. 38 No. 2, February 1991
- [11] Paper California Institut of Technology. *Minority Carrier Diffusion and Guard Structures*, CNS Memo No. 30 07/94
- [12] Texas Instruments Deutschland GmbH Learning Center. *Das Opto Kochbuch*, Technik Marketing München, 1. Auflage 1975
- [13] M. Loose. *Diplomarbeit über Layout und Test eines Systems adaptiver Photorezeptoren in analoger CMOS-Technologie*, Institut für Hochenergiephysik, Universität Heidelberg, 1996.
- [14] T. Delbrück, C. A. Mead. *Analog VLSI Phototransduction by continous-time, adaptive, logarithmic photoreceptor circuit*, Computation and Neural Systems Program, California Institute of Technology, Pasadena, Memo No. 30, 07/1994.



# Danksagung

Herzlichen Dank an alle, die zur Durchführung meiner Diplomarbeit beigetragen haben, insbesondere

Herrn Prof. K. Meier für die Bereitstellung dieses sehr interessanten Themengebietes und vorbildliche Betreuung,

Herrn Prof. U. Straumann für die Übernahme der Zweitkorrektur,

Herrn D. Droste ohne dessen fachkundige Beratung, Anregungen und die Bereitschaft *jederzeit* für Fragen zur Verfügung zu stehen, diese Arbeit so nicht zustande gekommen wäre, sowie für mühevoll und sorgfältiges Korrekturlesen des Manuskriptes,

Herrn M. Keller für die vielfältigen und freundlichen Hilfestellungen, aller Fragen bezüglich der Software-Infrastruktur jeglicher Art,

Herrn J. Schemmel für eingehende Diskussionen, Hilfsbereitschaft und die kritische und sorgfältige Durchsicht des Manuskriptes,

Herrn M. Loose für freundlichste Hilfe bei fachlichen sowie textgestaltenden Fragen,

Herrn J. Boelsems für die Durchführung diverser Meetings,

allen weiteren Mitgliedern des ASIC-Labors für die Hilfsbereitschaft und die stets angenehme Arbeitsatmosphäre,

allen Institutsmitarbeitern, die einen reibungsfreien technischen und organisatorischen Ablauf ermöglicht haben.



Erklärung:

Ich versichere, daß ich diese Arbeit selbständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 30.01.1998

.....